iejee
INTERNATIONAL ELECTRONIC
JOURNAL OF ELEMENTARY EDUCATION

www.iejee.com

December 2017 Volume 10 Issue 2

T&K AKADEMIC

iejee

INTERNATIONAL ELECTRONIC
JOURNAL OF ELEMENTARY EDUCATION

ISSN: 1307-9298

Editorial

Dear IEJEE Readers,

It is a great pleasure for me to present this issue of International Electronic Journal of Elementary Education. IEJEE by this issue is celebrating its 10th year. We are publishing four ordinary issues in 2017 in addition to a special issue.

In this number you'll find fourteen articles of which twenty-seven researchers stand behind from five different countries representing nineteen universities.

I would like to thank all of the contributors. Their research works encourage us as IEJEE editors.

I also would like to express my deepest gratitude to all the peer reviewers and the executive editors Dr. Gökhan Özsoy, Dr. Hayriye Gül Kuruyer and Dr. Hasan Tabak.

Editor-In-Chief
Dr. Kamil ÖZERK
University of Oslo

ISSN:1307-9298

All responsibility for statements made or opinions expressed in articles lies with the author.

Table of Content

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effect of Scaffolded Think-Group-Share Learning on Indonesian Elementary Schooler Satisfaction and Learning Achievement in English Classes

Octavia Mantik [a], Hee Jun Choi [b,*]

Received: 22 August 2017
Revised: 20 October 2017
Accepted: 25 October 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236113

Abstract

The purpose of this study was to examine whether or not "Scaffolded Think-Group-Share" learning can have a positive effect on student satisfaction and learning achievement in English classes of an Indonesian elementary school. To achieve this purpose, this study compared the findings from the two dependent variables (i.e., student satisfaction and English learning achievement) in "Scaffolded Think-Group-Share" learning with those in "Group Investigation" and "Learning Together" learning, which are other types of cooperative learning methods. According to the findings, there were statistically significant differences in student satisfaction and English learning achievement between the Scaffolded Think-Group-Share group, the Group Investigation group, and the Learning Together group. This study implies that "Scaffolded Think-Group-Share" learning has the potential to enhance student satisfaction and comprehension in English as a foreign language (EFL) classes.

Keywords: Cooperative learning, Scaffolding, Satisfaction, English learning achievement.

Introduction

The British Council (2013) reported that English is spoken by 1.75 billion people worldwide; a number that is projected to reach 2 billion by 2020. English has its status as a global language and has shown its dominance in some of the most important fields such as education, business, international relations, and politics (Crystal, 2003). Previous research found that the population's English skills are directly correlated with the country's economic performance (McCormick, 2013). This implies that proficiency in English is crucial for the welfare of individuals as well as for national development. As a result, non-English speaking countries invest enormous amounts of effort into prioritizing English education and emphasizing its role in globalization (Choi, 2007; Hu & McKay, 2012; Li, 2007; Tsui, 2004).

Meanwhile, Nunan (2003) found that many non-English speaking Asian countries have lowered the age for compulsory English education. China lowered the age from 11 to 9 in September 2011, and Korea lowered the age from 13 to 9 in 1995. In 2002, Taiwan introduced English as a compulsory subject for first graders, whereas in the past, English was first taught in the fifth grade. In Japan, English was not a compulsory subject in elementary school until April 2011, when the Ministry of Education, Culture, Science, and Technology launched the course of study for elementary schools—this required English to be compulsory starting in the fifth grade (Hu & McKay, 2012). In Indonesia, English has been introduced to fourth graders as a local content subject since 1993 (Rachmajanti, 2008). Generally, official English education begins at the primary school level in most non-English speaking Asian countries.

The Indonesian government, in particular, has taken much interest in improving teaching strategies in elementary school English classes in order to help children acquire English more effectively (Rachmajanti, 2008). According to previous studies conducted in Indonesia, many English teachers in Indonesian elementary schools employed a rather monotonous lecture method and seldom varied their teaching methods (Dardiri, 1994; Hawanti, 2011; Mardika, 2008; Rachmajanti, 2008; Sugeng, 2000; Zein, 2012). Rachmajanti (2008) revealed that most students in the Indonesian elementary schools were unsatisfied with the teacher-centered instructional method adopted by many Indonesian English teachers and preferred to work in small groups with interactive learning activities. These findings tend to be consistent with the contention of some scholars (e.g., Curtain & Dahlberg, 2016; Scott & Ytreberg, 1990) that young language learners (third, fourth, and fifth graders) work well in groups and learn from each other. Huda (1997) asserted that the monotonous teaching method in the Indonesian English classes was one of most critical obstacles to how well Indonesian children learn English.

Cooperative learning is regarded as one of instructional methods that can accommodate learning preferences of the Indonesian students and meet the need of the Indonesian government for improving teaching methods in elementary school English classes. Previous empirical

[a] Hongik University, Seoul, Korea . E-Mail: (O. Mantik) octaviamantik@gmail.com
[b,*] Corresponding author: Hee Jun CHOI, Associate Professor, Hongik University, 94 Wausan-ro, Mapo-gu, Seoul, 04066, South Korea, Phone: 82-2-320-1888, E-mail: hjchoi@hongik.ac.kr

studies revealed that cooperative learning is effective for enhancing learning achievement, developing higher-order thinking skills, encouraging pro-social behavior, improving inter-ethnic relationships, and increasing motivation to learn (e.g., Cohen, Lotan, & Catanzarite, 1990; Cook, 1984; High, 1993; Holt, 1993; Jacobs & Goh, 2007; Johnson, Johnson, & Stanne, 2000; McCafferty, Jacobs, & DaSilva Iddings, 2006; Sharan & Rich, 1984; Sharan & Shaulov, 1990; Slavin, 1995).

Previous research about cooperative learning in language classes showed that as a method, it supports reading, comprehension, and vocabulary development (e.g., Slavin, Lake, Chambers, Cheung, & Davis, 2009). Cooperative learning also tends to encourage more practice in language production (Deen, 1991). In particular, Alharbi (2008) investigated English reading performance among secondary school students aged 16 to 18 years old in Saudi Arabia and found a significant difference that favored cooperative learning compared to traditional teaching methods. Liao (2006) examined the impact of cooperative learning on the English grammar achievements of college students in Taiwan and found that cooperative learning had medium to large positive effects on grammar achievement. Nevertheless, cooperative learning is still considered neither as widely applied to young children nor systematically studied in English as foreign language (EFL) classrooms (Lan, Chang, & Sung, 2005; Lin, 2009; Ning, 2010). This implies that the effectiveness of cooperative learning in English as a foreign language needs to be more thoroughly investigated with learners of different ages, particularly students in elementary schools, by taking a closer look at the procedures of cooperative learning.

However, other research studies show that students who experienced cooperative learning do not always outperform their counterparts who received traditional instruction. Davidson (1985) conducted a thorough review of cooperative learning in mathematics education and found that one-third of the studies showed significant differences favoring cooperative learning versus traditional methods of instruction, whereas the remaining two-third of the studies did not show significant differences. A review study focusing on Asian students revealed that only half of the studies showed that cooperative learning had any positive effect on academic achievement (Thanh, Gillies, & Renshaw, 2008). Shaaban (2006) found no statistically significant difference between cooperative learning and whole-class instruction in improving English reading comprehension and vocabulary acquisition.

These inconsistent results of previous studies on the effectiveness of cooperative learning might be the result of differences in the main components of each cooperative learning method. Slavin (1990) argued that poorly constructed cooperative learning methods that lack the appropriate components can result in a free-rider effect, which is a major pitfall of cooperative learning. Thus, teachers should resolve the issue of free-rider effect in order to make cooperative learning more effective.

Some researchers assert that scaffolding can be designed to minimize the free-rider effect, which occurs when some group members do not their best because they assume that other group members will cover the work that they have to do (Janssen, Erkens, Kanselaar, & Jasper, 2006). According to Wood, Bruner, and Ross (1976), scaffolding refers to the "process that enables a child or novice to solve a problem, carry out a task or achieve a goal which would be beyond his unassisted efforts" (p. 90). The researchers went on to argue that this scaffolding primarily consists of receiving support from a teacher regarding the elements of the task that are initially beyond the child's ability, so that he or she can focus on the elements that are within his or her capacity. Support for scaffolding is needed more as the task becomes more complicated and the student's ability decreases (Donovan & Smolkin, 2002). In particular, most Indonesian elementary school students have not experienced cooperative learning and begin to learn English as a foreign language in the fourth grade. This implies that the cooperative learning that the Indonesian young children will receive needs to be integrated with well-structured scaffolding.

The concept of scaffolding is rooted in Vygotsky's social constructivist view of learning. Vygostky (1978) defined zone of proximal development as "the distance between the actual development level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). The definition of zone of proximal development implies that teaching is a process of co-construction of knowledge between the tutor and the learner and further transformation of that knowledge into individual knowledge of the learner (Verenikina, 2008). Vygostky (1978) argued that peer interaction and scaffolding are essential in facilitating individual cognitive growth and knowledge development. This means that cooperative learning that contains well-structured scaffolding can be an effective instructional method to help children construct knowledge and become self-regulated learners.

We developed a cooperative learning method that is termed "Scaffolded Think-Group-Share" learning and is implemented to resolve the problems of passive participation and task difficulty in cooperative learning in the English classes in an Indonesian elementary school based on the findings of previous studies on cooperative learning and scaffolding (e.g., Belland, Glazewski, & Richardson, 2008; McTighe & Lyman, 1988; Saye & Brush, 2002). Basically, it was developed based on the existing cooperative learning method called "Think-Pair-Share" that was developed by Frank Lyman in 1981, which includes three components: students think individually, talk with each other in pairs, and share their ideas with the larger group (McTighe & Lyman, 1988). Additionally, "Scaffolded Think-Group-Share" learning contains "hard scaffolding," which refers to static supports that are devised in advance based on typical student difficulties with a task (Saye & Brush, 2002) in the form of paper-based worksheet as a key component.

"Scaffolded Think-Group-Share" learning emphasizes individual activity prior to group activity by requiring students to work on a scaffolding worksheet individually to help them actively participate and cognitively engage in

group activity. "Scaffolded Think-Group-Share" learning might be able to help students enhance their learning achievement and satisfaction toward cooperative learning activity by minimizing the "free-riding" effect and promoting individual accountability. Accordingly, this study aimed to investigate the effect of "Scaffolded Think-Group-Share" learning on student satisfaction and learning achievement in English classes of an Indonesian elementary school.

More specifically, this study was intended to compare the findings from the two dependent variables (i.e., satisfaction and English learning achievement) in "Scaffolded Think-Group-Share" learning with those in "Group Investigation" and "Learning Together" learning, which are other types of cooperative learning methods in order to reveal which cooperative learning method is most effective in children's English classes. "Group Investigation" is a cooperative learning method in which the group task is divided among the group members; thus, each member does a unique part of the group task individually before synthesizing their answers as a group. "Learning Together" is another cooperative learning method in which students do group work without any preceding individual activity.

For the purposes of this study, the following research questions were addressed:

1. Does learner satisfaction in "Scaffolded Think-Group-Share" learning differ from those in "Group Investigation" and "Learning Together" learning?
2. Do learners' English learning achievements in "Scaffolded Think-Group-Share" learning differ from those in "Group Investigation" and "Learning Together" learning?.

Methodology

Research Design and Data Collection Procedure

This study used a three between-group, quasi-experimental design using a pretest and posttest. The details of the research design are as follows: Three fifth grade classes from an elementary school participated in this study. The three intact groups (already-formed classes) were randomly assigned to one of experimental, comparison, and control groups in the weeks that the treatments were given. Students in the experimental group were exposed to "Scaffolded Think-Group-Share" learning, students in the comparison group were exposed to "Group Investigation" learning, and students in the control group were exposed to "Learning Together" learning.

At the orientation session, all students in the three groups completed a pretest that covers all three lessons that were supposed to be taught during the experiment period. In addition, all the students in each group attended three lecture sessions from the same instructor on the first, third, and fifth week of the experiment period. The three posttests measured students' English learning achievements were administered as soon as the treatments were completed, in the second, fourth, and sixth week of the experiment period, in order to get information on the immediate learning outcomes related to the cooperative learning in which students participated. On the sixth week, students were also asked to express their satisfaction toward the cooperative learning methods that they experienced by filling-out a satisfaction questionnaire.

Table 1. *Three Between-Group Quasi-Experimental Design*

Cooperative learning	Orientation session	Right after the first, second, & third cooperative learning	Right after the third cooperative learning
X_{STGS}	O(Pre)	O($Post_1$, $Post_2$, $Post_3$)	O(S)
X_{GI}	O(Pre)	O($Post_1$, $Post_2$, $Post_3$)	O(S)
X_{LT}	O(Pre)	O($Post_1$, $Post_2$, $Post_3$)	O(S)

Note. X_{STGS} = "Scaffolded Think-Group-Share", X_{GI} = "Group Investigation", X_{LT}= "Learning Together", O= Test or Survey, (Pre)= Pretest of the first, second, and third lessons ($Post_1$)= Comprehension test of the first lesson, ($Post_2$)= Comprehension test of the second lesson, ($Post_3$)= Comprehension test of the third lesson, (S)= Satisfaction.

Participants

This research was conducted in an elementary school near Jakarta, Indonesia. Participants were 111 fifth grade students who belonged to one of the three classes. Of the 111 participants, 55 students (49.5 %) were male, and 56 students (50.5 %) were female. Each class was randomly assigned into one of the three groups (experimental group, comparison group, or control group). The numbers of the students in the experimental, comparison, and control groups were 38, 36, and 37, respectively. The three groups were taught by one English teacher who willingly agreed to participate in this study. He holds qualification as an English teacher and has taught English in the elementary school since 2001. The English teacher delivered a lecture to the students in the three groups in the week before the treatments were given and administered the treatments for each group according to the lesson plans developed for this study.

Instrumentation

This study employed a survey questionnaire with a five-point symmetrical Likert scale that consisted of six items in order to measure students' satisfaction toward cooperative learning activities that they experienced in their groups. The questionnaire was adapted from the satisfaction subscale in the Instructional Materials Motivation Survey (IMMS) originally developed by Keller (1987). Previous research studies showed that this instrument had a Cronbach's alpha of .86 (Choi & Johnson, 2005) and .89 (Choi & Johnson, 2007).

Additionally, this study employed one pretest and three posttests that focused on the learning objectives of the English lessons for which the treatments were provided.

The pretest consisted of thirty questions to measure students' prior knowledge of the contents of three lesson units to be taught during the experiment. Each posttest consisted of ten questions to measure students' immediate learning outcomes related to the cooperative learning that they received for each lesson unit. Both the pretest and posttests were divided into two sections: multiple choice and short answer questions. The questions in each posttest were identical to those in the pretest; however, the sequence of the numbers and options in the multiple-choice questions were changed randomly. The tests were developed by an English teacher who has considerable experience in setting questions for an English test. Finally, two professors reviewed the tests to check whether the content validity can be ensured.

Treatments

The treatments were given to the students in the span of six weeks. On the first, third, and fifth week of the experiment period, the students in each group (experimental, comparison, or control group) attended lecture sessions from the same English teacher. Each lecture session covered different units of English instruction. On the second, fourth, and sixth week, the students in each group participated in a different type of cooperative learning: "Scaffolded Think-Group-Share" learning (experimental treatment), "Group Investigation" learning (comparison treatment), or "Learning Together" learning (control treatment).

In "Scaffolded Think-Group-Share" cooperative learning, students progressed through three steps. In the "Scaffolded Think" component, students were asked to work on the scaffolding worksheet individually. This scaffolding worksheet was a tool designed to help students activate prerequisite knowledge and develop certain skills needed to complete the group task. The scaffolding worksheets included cues and questions that can prompt students to think about certain concepts, which are relevant to solving the corresponding group tasks. Shortly after, students continued to the "Group" component, in which they worked together with other group members to complete the group task that is related to the lesson taught by the English teacher in the previous week. In the final component, the "Share" component, groups shared their answers with the rest of the class.

The comparison group was exposed to "Group Investigation" cooperative learning, in which students were asked to divide the group task. For this activity, each group member was given a unique task to finish individually within a certain amount of time. This is also called task specialization, a means by which, according to Slavin (1983), individual accountability can be established. Each group synthesized all group members' work to complete the group task and help each other master the materials. Lastly, groups presented their answers in front of the whole class. Meanwhile, the control group was exposed to "Learning Together" cooperative learning, which is identical to the second and third components—namely, the "Group" and "Share" components—of the "Scaffolded Think-Group-Share" cooperative learning. The control group did not initiate individual activity prior to group activity.

The three classes (experimental, comparison, and control groups) had small learning groups that consisted of three or four students who were engaged in cooperative learning. The students in each class were randomly assigned to the small learning groups. Each student in the small learning group took on one of the following roles: moderator, timekeeper, note-taker, or presenter. Each group's moderator was responsible for stimulating group members to elicit their ideas, finalizing group discussion, and making sure that each group member understood the material. The timekeeper was responsible for ensuring that the group could finish the task within the given time. The note-taker's responsibilities were to take notes about group discussions and summarize explanations from the teacher. The presenter was responsible for presenting the group's answers, and answering questions from peers in other groups and the teacher as a group's spokesperson. In the small group consisting of three students, the moderator also assumed the timekeeper's role.

Data Analysis

The quantitative data used to answer the first research question were analyzed using one-way Analysis of Variance (ANOVA), which determines whether there is a statistically significant mean difference in a dependent variable between two or more groups with one independent variable (Gall, Gall, & Borg, 2007). Additionally, this study employed Multivariate Analysis of Covariance (MANCOVA) as a statistical technique to answer the second research question. MANCOVA was used to determine whether there were significant mean differences in two or more measured variables (i.e., comprehension of the first lesson, comprehension of the second lesson, and comprehension of the third lesson) that were correlated between groups while controlling for the confounding factor (Hair, Black, Babin, & Anderson, 2010).

Results

Differences in Learner Satisfaction between the Three Groups

One-way ANOVA was conducted to investigate whether learner satisfaction in "Scaffolded Think-Group-Share" learning differs from those in "Group Investigation" and "Learning Together" learning. The independent variable represented the different types of cooperative learning methods with three groups (i.e., the Scaffolded Think-Group-Share group, the Group Investigation group, and the Learning Together group). The dependent variable was the mean score that students made on a survey questionnaire that was used to gauge satisfaction toward cooperative learning activities that they experienced in their groups. The questionnaire consisted of six items scored using a five-point Likert scale. Table 2 presents the means and standard deviations of learner satisfaction for each of the three groups.

Table 2. *Means and Standard Deviations of Learner Satisfaction by Group*

Group	*n*	*M*	sd
Scaffolded Think-Group-Share	38	29.00	1.49
Group Investigation	36	23.42	3.51
Learning Together	37	22.78	3.41
Total	111	25.12	4.06

Note. Maximum high score in learner satisfaction= 30.

The normality of the learner satisfaction variable was tested in terms of its skewness and kurtosis before further analyses could proceed. West, Finch, and Curran (1995) argue that the criteria of normality are skewness <2 and kurtosis <7. The normality assumption of learner satisfaction was met because its skewness and kurtosis were - 0.58 and 0.08, respectively. However, the *Levene's F* test indicated that the homogeneity of variance assumption was not satisfied at the 0.05 level. Thus, the *Welch's F* test was employed. The results of one-way ANOVA revealed that there were statistically significant differences in learner satisfaction between the Scaffolded Think-Group-Share group, the Group Investigation group, and the Learning Together group [*Welch's F*(2, 60.05)= 78.03, p<.001] .

Post hoc comparisons were conducted using the Games-Howell post hoc procedure because the homogeneity of variance assumption was not met. The results presented in Table 3 indicate that students in the Scaffolded Think-Group-Share group (*M*= 29.00, sd= 1.49) had a significantly higher mean score for learner satisfaction than students in the Group Investigation group (*M*= 23.42, sd= 3.51) as well as students in the Learning Together group (*M*= 22.78, sd= 3.41) at the .001 level. The effect sizes for the two significant effects were 2.07 and 2.36, respectively.

Table 3. *Post Hoc Results for Learner Satisfaction by Group*

Group	Mean difference (Effect size)		
	1	2	3
Scaffolded Think-Group-Share	-		
Group Investigation	- 5.58*** (2.07)	-	
Learning Together	- 6.22*** (2.36)	- 0.63	-

***p < .001.

Differences in Learners' English Learning Achievements between the Three Groups

MANCOVA was employed to determine whether the type of cooperative learning (i.e., Scaffolded Think-Group-Share, Group Investigation, and Learning Together) affected students' English learning achievements while controlling for their pretest scores as the covariate. Table 4 shows the means and standard deviations of English learning achievement in the three comprehension tests for each of the three groups.

Table 4. *Means and Standard Deviations in the Three Comprehension Tests by Group*

Variables	Group	*n*	*M*	sd
Comprehension 1	Scaffolded Think-Group-Share	38	41.32	19.75
	Group Investigation	36	16.39	15.34
	Learning Together	37	15.41	15.02
	Total	111	24.59	20.66
Comprehension 2	Scaffolded Think-Group-Share	38	65.79	15.53
	Group Investigation	36	41.39	15.52
	Learning Together	37	22.16	14.36
	Total	111	43.33	23.48
Comprehension 3	Scaffolded Think-Group-Share	38	49.74	19.93
	Group Investigation	36	31.39	18.03
	Learning Together	37	21.35	19.28
	Total	111	34.32	22.37

Note. Maximum high score in each comprehension test = 100.

The test for normality indicated that the data from each comprehension test were statistically normal. The *Levene's F* test indicated that the homogeneity of variance assumption for each dependent variable (comprehension 1, 2, and 3) was met at the .05 level (p= .36, p= .62, and p= .70, respectively). Additionally, the results of the Box's M test showed that the assumption of covariance equality between the three groups was met at the .05 level (p= .25). Accordingly, we conducted a MANCOVA analysis. As shown in Table 5, the MANCOVA results indicate that there was a statistically significant difference in English learning achievements between students who participated in Scaffolded Think-Group-Share learning, students who participated in Group Investigation learning, and students who participated in Learning Together (Wilks' lambda=

0.38, p<.001), with the effects of prior knowledge (Wilks' lambda= 0.78, p<.001) being controlled.

More specifically, students who participated in the Scaffolded Think-Group-Share learning showed significantly higher English learning achievement than students who participated in the Group Investigation learning and students who participated in the Learning Together learning in the first and second comprehension tests at the .001 level. In the third comprehension test, students who participated in the Scaffolded Think-Group-Share learning significantly outperformed students who participated in the Group Investigation learning at the .01 level while they significantly outperformed students who participated in the Learning Together learning at the .001 level. Table 6 indicates the mean differences in the three comprehension tests between the Scaffolded Think-Group-Share learning, Group Investigation learning, and Learning Together learning groups and the effect sizes.

Table 5. *MANCOVA Results: Multivariate Tests*

Variables	Wilks' lambda	*F*	*p*
Prior knowledge	0.78	9.63	0.000***
Comprehension 1	0.38	21.59	0.000***
Comprehension 2			
Comprehension 3			

***p < .001.

Table 6. *Contrast Analysis Results for English Learning Achievement by Group*

Variables	Group	Mean difference (Effect size)		
		1	2	3
Comprehension 1	Scaffolded Think-Group-Share	-		
	Group Investigation	- 22.36***	-	
		(1.41)		
	Learning Together	- 21.43***	0.94	-
		(1.48)		
Comprehension 2	Scaffolded Think-Group-Share	-		
	Group Investigation	- 24.77***	-	
		(1.57)		
	Learning Together	- 44.28***	- 19.51***	-
		(2.92)	(1.29)	
Comprehension 3	Scaffolded Think-Group-Share	-		
	Group Investigation	- 14.70**	-	
		(0.97)		
	Learning Together	- 22.02***	- 7.31	-
		(1.45)		

***p <.001, **p <.01.

Conclusions and Discussion

According to this study, "Scaffolded Think-Group-Share" learning is a more effective cooperative learning method for learner satisfaction than "Group Investigation" and "Learning Together" learning are in English classes for young children. This result might be supported by previous findings that indicate that college students felt more satisfied with "Think-Pair-Share" learning including a procedural scaffolding, which led to more active student participation in the group work, than other types of instructional methods (Lange, Costley, & Han, 2016).

Because "Scaffolded Think-Group-Share" learning was developed based on "Think-Pair-Share" learning and contains hard scaffolding, it might help students activate prerequisite knowledge effectively and develop certain skills needed to complete the group task. In other words, "Scaffolded Think-Group-Share" learning contains much more elaborate and structured scaffolds than "Think-Pair-Share" learning, which contains a basic scaffolding, as well as "Group Investigation" and "Learning Together," which almost never contains scaffolding components. In addition, the empirical study, which was conducted by Saye and Brush (2002), revealed that hard scaffolding was helpful in decreasing learners' cognitive loads and making them engage in learning activities more actively. This implies that hard scaffolding can play a critical role in enhancing learner satisfaction in a cooperative learning environment. Consequently, individual activity on the scaffolding worksheet conducted in "Scaffolded Think-Group-Share" learning might have positively affected how young students had more increased satisfaction toward their cooperative learning activities in English classes by inducing more active student participation.

This study also shows that "Scaffolded Think-Group-Share" learning is more effective than "Group Investigation" and "Learning Together" learning for enhancing learner comprehension in English classes for young children. This result is congruent with the findings of previous studies. Some researchers showed that scaffolding had a positive influence on students' learning outcomes or performance in a collaborative or problem-based learning context (Huang, Wu, & Chen, 2012; Simons & Klein, 2007). Additionally, Pea (2004) contended that well-designed scaffoldings can help students successfully resolve and implement complex problems or tasks.

"Scaffolded Think-Group-Share" learning included a well-structured scaffolding component consisting of cues and questions that prompted students to think about certain concepts that were relevant for solving the corresponding group tasks. This scaffolding component of "Scaffolded Think-Group-Share" learning might have helped students cognitively engage in the group task by activating prerequisite knowledge individually prior to group activity. This might be the crucial reason why students in "Scaffolded Think-Group-Share" group outperformed their counterparts in "Group Investigation" and "Learning Together" groups in the three comprehension tests.

This study might be significant in that it was an initial effort to determine the actual impact of individual activity using hard scaffolds in cooperative learning. Additionally, there are few empirical studies on cooperative learning that have been conducted with young children in EFL classrooms (Lan, Chang, & Sung, 2005; Lin, 2009; Ning, 2010). Consequently, the findings of this study contribute to expanding and fortifying the existing knowledge base regarding cooperative learning using well-structured scaffolding.

Limitations and Suggestions for Future Research

This study has the following limitations, despite its significance. First, the findings of this study may not be generalized to all elementary students because the sample was selected only from those students who attend an elementary school near Jakarta, Indonesia. Accordingly, future research studies should be implemented that use a more extensive and larger sample in terms of regional, ethnical, and cultural backgrounds so that the findings can be generalized to all elementary students. Second, the findings of this study were drawn from only quantitative data, which provides limited information regarding the relative advantages of "Scaffolded Think-Group-Share" learning. Future research studies need to produce findings using both quantitative and qualitative data in order to provide in-depth information on "Scaffolded Think-Group-Share" learning.

Acknowledgement:

This work was supported by the Hongik University Research Fund.

References

Alharbi, L. A. (2008). *The effectiveness of using cooperative learning method on ESL reading comprehension performance, students' attitudes toward CL, and students' motivation toward reading of secondary stage in Saudi public girls' schools* (Doctoral dissertation, West Virginia University). Retrieved from http://search.proquest.com/docview/304448780

Belland, B., Glazewski, K., & Richardson, J. (2008). A scaffolding framework to support the construction of evidence-based arguments among middle school students. *Educational Technology Research and Development*, *56*(4), 401-422.

British Council. (2013). *The English effect*. Retrieved from https://www.britishcouncil.org/sites/default/files/english-effect-report-v2.pdf

Choi, H., & Johnson, S. D. (2005). The effect of context-based video instruction on learning and motivation in online courses. *The American Journal of Distance Education, 19*(4), 215-227.

Choi, H., & Johnson, S. D. (2007). The effect of problem-based video instruction on learner satisfaction, comprehension and retention in college courses. *British Journal of Educational Technology, 38*(5), 885-895.

Choi, Y. H. (2007). The history of and the policy of the English language education in Korea. In Y. H. Choi & B. Spolsky (Eds.), *English education in Asia: History and policies* (pp. 33-66). Seoul: Asia TEFL.

Cohen, E. G., Lotan, R. A., & Catanzarite, L. (1990). Treating status problems in the cooperative classroom. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 173-202). New York: Praeger.

Cook, S. (1984). Cooperative interaction in multiethnic contexts. In N. Miller & M. Brewer (Eds.), *Groups in contact: The psychology of desegregation* (pp. 77-96). New York: Academic.

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge: Cambridge University Press.

Curtain, H., & Dahlberg, C. A. (2016) *Languages and learners: Making the match*. Boston: Pearson.

Dardiri, H. (1994). *A study on the teaching of English at some elementary schools in Kodya Malang* (Unpublished thesis). IKIP Malang, Indonesia.

Davidson, N. (1985). Small group learning and teaching in mathematics: A selective review of the research. In R. Slavin, S. Sharan, S. Kagan, R. Hertz-Lazarowitz, G. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn*. New York: Plenum.

Deen, J. Y. (1991). Comparing interaction in a cooperative learning and teacher-centered language classroom. *I. T. L. Review of Applied Linguistics*, 153-181.

Donovan, C. A., & Smolkin, L. B. (2002). Children's genre knowledge: An examination of K-5 students' performance on multiple tasks providing differing levels of scaffolding. *Reading Research Quarterly*, *37*, 428-465.

Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: An introduction* (8th ed.). Berkshire: Allyn and Bacon.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Upper Saddle River, NJ: Pearson.

Hawanti, S. (2011). Teaching English in Indonesian primary schools: The missing link. *Leksika*, *5*(1), 62-69.

High, J. (1993). *Second language learning through cooperative learning*. San Juan Capistrano, CA: Kagan Cooperative Learning.

Holt, D. (Ed.). (1993). *Cooperative learning: A response to linguistic and cultural diversity*. McHenry, IL: Centre for Applied Linguistics.

Hu, G. W., & McKay, S. L. (2012). English language education in East Asia: Some recent developments. *Journal of Multilingual and Multicultural Development, 33*(4), 345-362.

Huang, H., Wu, C., & Chen, N. (2012). The effectiveness of using procedural scaffoldings in a paper-plus-smartphone collaborative learning context. *Computers & Education, 59*(2), 250-259.

Huda, N. (1994). The teaching of English in primary schools: Issues and problems. *TEFLIN Journal*, *6*(2), 82-90.

Jacobs, G. M., & Goh, C. C. (2007). *Cooperative learning in the language classroom*. Singapore: SEAMEO Regional Language Centre.

Janssen, J., Erkens, G., Kanselaar, G., & Jasper, J. (2006). Visualization of participation: does it contribute to successful computer-supported collaborative learning? *Computers & Education, 49*(4), 1037-1065.

Johnson, D. W., Johnson, R. T., & Stanne, M. E. (2000). *Cooperative learning methods: A meta-analysis*. Minneapolis, MN: University of Minnesota Press.

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, *10*(3), 2-10.

Lan, Y. J., Chang, K. E., & Sung, Y. T. (2005). Cooperative learning in small EFL early reading groups. *Journal of Pan-Pacific Association of Applied Linguistics*, *9*(2), 319-341.

Lange, C., Costley, J., & Han, S. L. (2016). Informal cooperative learning in small groups: The effect of scaffolding on participation. *Issues in Educational Research, 26*(2), 260-279.

Li, M. (2007). Foreign language education in primary schools in the People's Republic of China. *Current Issues in Language Planning*, *8*(2), 148-160. doi: 10.2167/cilp113.0

Liao, H. C. (2006). *Effects of Cooperative Learning on Motivation, Learning Strategy Utilization, and Grammar Achievement on English language Learners in Taiwan* (Doctoral dissertation, University of New Orleans). Retrieved from http://scholarworks.uno.edu/cgi/viewcontent.cgi?article=1362&context=td

Lin, M. (2009). *Effects of cooperative learning on the oral proficiency of Chinese students in the tertiary-level EFL classroom* (Doctoral dissertation, University of Leicester). Retrieved from https://lra.le.ac.uk/bitstream/2381/8941/1/2009linmxedd.pdf

Lyman, F. (1981). The responsive classroom discussion. In A. S. Anderson (Ed.), *Mainstreaming Digest* (pp. 109-113). College Park, MD: University of Maryland College of Education.

Mardika, I. N. (2008). *Pengembangan multimedia dalam pembelajaran kosakata Bahasa Inggris di SD* [The development of multimedia for English vocabulary learning in elementary school] (Unpublished thesis). Yogyakarta State University, Indonesia.

McCafferty, S. G., Jacobs, G. M., & DaSilva Iddings, A. C. (Eds.). (2006). *Cooperative learning and second language teaching*. Cambridge: Cambridge University Press.

McCormick, C. (2013, November 15). Countries with better English have better economies. *Harvard Business Review*. Retrieved from https://hbr.org/2013/11/countries-with-better-english-have-better-economies

McTighe, J. & Lyman, F. T. (1988). Cueing thinking in the classroom: The promise of theory-embedded tools. *Educational Leadership, 45*(7), 18-24.

Ning, H. (2010). *An investigation of the use of cooperative learning in teaching English as a foreign language with tertiary education learners in China* (Doctoral dissertation, University of Canterbury). Retrieved from https://ir.canterbury.ac.nz/bitstream/handle/10092/5188/thesis_fulltext.pdf?sequence=1&isAllowed=y

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly*, *37*(4), 589-613.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences, 13*(3), 423-451.

Rachmajanti, S. (2008). Impact of English instruction at the elementary schools on the students' achievement of English at the lower secondary school. *TEFLIN Journal*, *19*(2), 160-185.

Saye, J. W., & Brush, T. (2002). Scaffolding critical reasoning about history and social issues in multimedia-supported learning environments. *Educational Technology Research and Development*, *50*(3), 77-96.

Scott, W. A., & Ytreberg, L. H. (1990). *Teaching English to children*. New York: Longman.

Shaaban, K. (2006). An initial study of the effects of cooperative learning on reading comprehension, vocabulary acquisition, and motivation to read. *Reading Psychology*, *27*(5), 377-403. doi: 10.1080/02702710600846613

Sharan, S., & Rich, Y. (1984). Field experiments on ethnic integration in Israeli schools. In Y. Amir & S. Sharan (Eds.), *School desegregation* (pp. 189-217). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Sharan, S., & Shaulov, A. (1990). Cooperative learning, motivation to learn, and academic achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 173-202). New York: Praeger.

Slavin, R. E. (1983). *Cooperative learning*. New York: Longman.

Slavin, R. E. (1990). *Cooperative learning: Theory, research and practice*. Englewood Cliffs, NJ: Prentice Hall.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (2nd ed.), Boston: Allyn and Bacon.

Slavin, R. E., Lake, C., Chambers, B., Cheung, A., & Davis, S. (2009). Effective reading programs for the elementary

grades: A best-evidence synthesis. *Review of Educational Research*, *79*, 1391-1466.

Sugeng, B. (2000). Pembelajaran Bahasa Inggris untuk anak: Sifat-sifat pembelajaran yang baik [English learning for children: Good learning characteristics] (Unpublished paper). Yogyakarta State University, Indonesia.

Thanh, P. T. H., Gillies, R., & Renshaw, P. (2008). Cooperative learning and academic achievement of Asian students: A true story. *International Education Studies*, *1*, 83-88.

Tsui, A. B. M. (2004). Language policies in Asian countries: Issues and tensions. *The Journal of Asia TEFL*, *1*, 1-25.

Verenikina, I. (2008). Scaffolding and learning: Its role in nurturing new learners. In P. Kell, W. Vialle, D. Konza & G. Vogl (Eds.), *Learning and the learner: Exploring learning for new times* (pp. 161-180). Wollongong, Australia: University of Wollongong.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

West, S.G., Finch, J.F. & Curran, P.J. (1995). Structural equation models with nonnormal variables. Problems and remedies. In R.H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues and applications* (pp. 56-75). Newbury Park, CA: Sage.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *17*, 89-100.

Zein, M. S. (2012). *Language teacher education for primary school English teachers in Indonesia: Policy recommendations* (Doctoral dissertation, Australian National University). Retrieved from https://open research-repository.anu.edu.au/bitstream/1885/9981/1/Zein_M.S._2012.pdf

This page is intentionally left blank

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

Adaptation of Teachers' Conceptions and Practices of Formative Assessment Scale into Turkish Culture and a Structural Equation Modeling*

Pınar Karaman [a,**], Çavuş Şahin [b]

Received: 21 October 2017
Revised: 28 November 2017
Accepted: 30 November 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236114

Abstract

The purpose of this study was to adapt Teachers' Conceptions and Practices of Formative Assessment Scale (TCPFS) based on the Theory of Planned Behavior (TPB) into Turkish culture and apply the TPB to examine teachers' intentions and behaviors regarding formative assessment. After examining linguistic validity of the scale, Turkish scale was applied to 400 primary school teachers in Turkey. Confirmatory factor analysis (CFA) was administered for the construct validity. The results showed that the model had fairly good model fit (RMSEA=.07, CFI= .90, TLI=.87, IFI= .90). Reliability of the Turkish version of the scale was attained through Cronbach's alpha coefficient. Findings revealed that adapted scale has satisfactory psychometric features. Since the psychometric properties of the scale in the TPB were acceptable, predictive model was estimated. Structural equation modeling was used to examine the applicability of TPB in understanding teachers' conceptions and practices of formative assessment. The results supported that TPB model can help to predict and explain teachers' conceptions and behaviors regarding formative assessment.

Keywords: Formative assessment; teachers; theory of planned behavior.

Introduction

Today, assessment is not only used for accountability purposes but also used for shaping ongoing instruction in the classroom (Darling-Hammond & Pecheone, 2010; Gong, 2010; Wilson & Draney, 2004). This assessment approach called as assessment for learning approach or formative assessment is intended to give information about student learning (Black & William, 2006; Heritage, 2010a; Shepard, 2000). Therefore, it distinguishes from diagnostic assessment approach used to identify students and summative assessment approach used to make final judgment of student competency (Trumbull & Lash, 2013). Formative assessment approach based on cognitive theory and sociocultural theory provides several activities: giving continuous feedback, promoting students' participation, making instructional changes in the classroom (Black & Wiliam, 1998; NCTE, 2013; Shermis & Di Vesta, 2011). Since this type of assessment is to shape or help students' learning during the learning process, it is called as a process during ongoing instruction (Black & William, 1998; Frohbeiter, Greenwald, Stecher, & Schwartz, 2011; Sadler, 1998). Heritage (2007) also defined formative assessment as in four core elements: facilitating student learning, meaningful feedback, student involvement, and learning progressions. Many studies addressed that formative assessment practices have positive effect on students' learning and attainment (Bennett, 2011; Hattie, 2009; Heritage, 2010b; Ruiz-Primo & Furtak, 2006; Sadler, 1998; Sumantri & Satriani, 2016; Jonsson, Lundahl, & Holmgren, 2015; Volante & Beckett, 2011). Therefore, assessment for learning as part of the education reform has been largely focused in many countries such as UK, Australia, Hong Kong, New Zealand, USA etc. The new assessment culture has also been placed in Turkey's current curriculum (MoNE, 2013).

However implementation of educational reforms in schools is a challenging and complex process (Fullan, 1999). Teachers have a major role for achieving educational change (Carles, 2015; Knight, 2002). Teachers' educational philosophies and their conceptions consisted of their beliefs, attitudes, and intentions may have important effect on their behaviors in the classroom (Brown, 2004; Pajares, 1992; Harrison, 2013; Haney & McArthur, 2002; Woolfolk Hoy, Davis, & Pape, 2006). Specifically the quality of assessment for learning approach or formative assessment implementation may rely heavily on beliefs, attitudes, and intentions that teachers have (Brown, Harris, & Harnett, 2012). Several research has been done about teachers' beliefs, attitudes, and practices regarding formative assessment (Carles, 2015; Brown & Gao, 2015; Brown, Kennedy, Fok, Chan, & Yu, 2009; Davis & Neitzel, 2011; Young & Jackman, 2014; Yau, 2004). In Turkey, the research about teachers' formative assessment beliefs, attitudes or practices were very limited (Aydoğmuş & Keskin, 2012; Konur & Konur, 2011; Öz, 2014; Sönmez Ektem & Erben Keçici, & Pilten, 2016). Most of the studies related with teachers' attitudes, beliefs or practices regarding formative assessment focused on only one aspect of assessment but relationship between these variables under a theoretical framework has not been extensively investigated (Yan & Cheng, 2015).

*This study was orally presented at the Ninth International Congress of Educational Research on 11-14 May, 2017 held in Ordu University, Turkey

[a,*] Corresponding author: Pınar KARAMAN, Sinop University, Faculty of Education, Department of Educational Sciences, Sinop/Turkey. E-mail: pkaraman1626@gmail.com

[b] Çanakkale Onsekiz Mart University, Faculty of Education, Department of Primary Education, Çanakkale/Turkey. E-mail: csahin25240@yahoo.com

Theory of planned behavior

Theory of planned behavior (TPB) as an explanatory model proposed by Ajzen (1985, 1991) explains human behavior's intention to perform actual behavior (Figure 1). The theory gives the relationships among attitude, subjective norms, perceived behavioral control, intention, and behavior. According to the theory, a person with more favorable attitude, subjective norm, and perceived behavioral control will more likely to have intention to perform behavior. Bandura's (1997) self-efficacy theory used as an indicator of perceived difficulty also has been used in TPB model (e.g. Martin & Kulinna, 2004; Tery & O'Leary, 1995; Yan & Cheng, 2015).

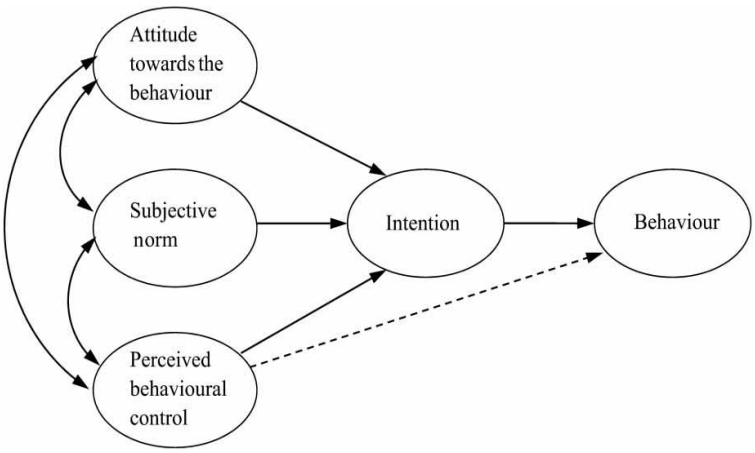


Figure 1. Theory of planned behavior (Ajzen, 1991)

TPB suggests that four factors impact on individuals' behaviors: (a) intentions, (b) attitudes, (c) subjective norm, (d) perceived behavioral control has both an indirect (via intentions) and direct impact.

The power of TPB model to predict and understand human behaviors was examined in many fields (e.g., Conner, Warren, Close, Sparks, 1999; Conatser, Block, & Gansneder, 2002; Huchting, Lac, & LaBrie, 2008; Martin & Kulinna, 2005; Schifter & Ajzen, 1985). TPB has also been used to understand teachers' attitudes, intentions, and behaviors (Crawley, 1990; Zint, 2002; Kersaint, Lewis, Potter, & Meisels, 2007; Stanec, 2009). However, there are a few studies which used TPB to investigate teachers' conceptions and behaviors regarding assessment (e.g., Yan, 2014; Yan & Cheng, 2015; Yim & Cho, 2016). Since very limited studies have specifically investigated teachers' formative assessment conceptions and practices in the TPB model (e.g., Yan & Cheng, 2015), more empirical research in that field is important.

The purpose of study was twofold. Firstly, adaptation of the Teachers' Conceptions and Practices of Formative Assessment Scale (TCPFS) developed by Yan and Cheng (2015) to Turkish was conducted in the study. Since the TCPFS scale based on TPB model provided specifically the prediction and explanation of teachers' intentions and practices regarding formative assessment, adaptation of the scale into Turkish culture is very important. With adapted scale, Turkish teachers' attitudes, intentions, and practices regarding formative assessment under a theoretical framework will be investigated extensively in empirical research. Therefore, the TCPFS scale in Turkish version will contribute to the relevant literature in Turkey. In the second part of the study, applicability of TPB was examined to understand the relationships among teachers' attitudes, intentions, and practices of formative assessment.

Methodology

Adaptation of an instrument is a complex task. The process of translation, adaptation and validation of an instrument for use in other cultures and languages requires rigorous planning and most appropriate methodology (Bracken & Barona, 1991; Brislin, 1970). In the first phase of the study, validity and reliability of Turkish version of TCPS scale was examined. In the second phase of the study, predictive model was examined by using structural equation modeling (SEM).

Participants

In this study, "Teachers' Conceptions and Practices of Formative Assessment Scale (TCPFS)" was carefully adapted and tested to ensure its relevance to teachers in Turkey. The sample consisted of 400 primary teachers working at public schools in the province of Çanakkale and Sinop in Turkey.

Instrument

TCPFS has been developed by Yan and Cheng (2005) for measuring primary teachers' conceptions and practices of formative assessment. The questionnaire consisted of 40 items and seven scales in the TPB model regarding formative assessment. The instrument had affective attitude scale (AAT) with 7 items, instrumental attitude scale (IAT) with 13 items, subjective norm scale (SNO) with 5 items, controllability scale (CON) with 4 items, self-efficacy scale (SEF) with 6 items, intention scale (INT) with 6 items, and behavior scale (BEH) with 2 items. Rasch analysis and path analysis were used for examining the psychometric properties of the scale. The result of the analyses indicated a satisfactory fit between proposed model and observed data (Explained variance: 51%, chi-square statistic $\chi 2$= 7.678, p>0.05; RMSEA= .059; GFI= .995; CFI= .997; TLI= .971). Therefore, the instrument developed by Yan and Cheng (2005) using TPB as a

theoretical framework regarding formative assessment was psychometrically robust enough.

Procedure

The adaptation of the TCPF scale into Turkish was done in accordance with Hambleton's (2005) suggestions. Hambleton's (2005) suggestions about five essential stages for instrument adaptation were considered in this study:

1. Original instrument was translated into Turkish by two bilingual translators who are familiar with the subject. These two translators produced the initial translations independently.

2. After the two versions of the translated instrument by two bilingual translators, these translations were compared and examined. With this examination, translated versions were compared with the original instrument in terms of their semantic, idiomatic, experiential, and conceptual equivalence. At the end of this process, researchers attained one single version of translated instrument.

3. The translated instrument was presented to the group of teachers (*N*=10) to verify whether the items of the instrument were understandable for the intended group. The teachers expressed that the items of the instrument were clear and understandable.

4. Back translation was used. The translated version of the instrument was translated back into the source language by two translators other than first ones. Back translated version of the instrument and the original instrument were compared. The result showed that two versions were conceptually similar.

5. The instrument was piloted with a small group of teachers (*N*=50). The instructions and items of the translated instrument were examined in this process. Pilot study result showed that the translated instrument was ready for the teachers as a target group. After administering the instrument to the participants, the psychometric properties of the instrument was examined by using recognized statistical analysis.

Findings

Construct Validity

In this study, confirmatory factor analysis was used to assess how well hypothesized measurement model data fit with observed data (Byrne, 2001; Tabachnick & Fidell, 2013). In accordance with this aim, AMOS program was used for CFA. Since some items' factor loading values were less than .40, these items were eliminated from the scale (AAT7 item, IAT1 item, and SEF1item). After eliminating these items from the scale, CFA was runned again. The results of CFA fit indices showed the sufficiency of the model (χ^2= 2301,272, *df*=609 *p*<.001, RMSEA=.08, CFI= .86, TLI=.85, IFI= .86). For RMSEA, .01, .05, and .08 values indicate excellent, good, and acceptable fit (MacCallum, Browne, & Sugawara, 1996; Sümer, 2000). Therefore, the value of .08 for RMSEA was considered as acceptable fit. There are several goodness of fit indices for CFA such as CFI, TLI, NFI, IFI. The value of .90 was considered as adequate fit. And the value of .95 was considered as a perfect fit (Bentler & Bonett, 1980; Şimşek, 2007). The goodness of fit indices in this study indicated the acceptable model fit. Besides, CFA suggested some modifications between some items. Some residuals were allowed to correlate since the variables had on the same factor (Landis, Edwards, & Cortina, 2009). After the modifications between some items, the adjusted model showed that the model had fairly good model fit (χ^2=1906.27, *df*=567, *p*<.001, RMSEA=.07, CFI= .90, TLI=.87, IFI= .90). Results from CFA are presented in Figure 2.

Reliability

Reliability of the Turkish version of the TCPFS scale was attained through Cronbach's Alpha coefficient (Table 1). Internal consistency of the scale were found .95 for the "whole scale", .90 for AAT sub-factor, .93 for IAT sub-scale, .83 for SN sub-scale, .92 for CON sub-scale, .86 for SEF sub-scale, .93 for INT sub-scale and .86 for BEH sub-scale. The results showed that overall total scores and also sub-scales had good reliability.

Table 1. *Cronbach's Alpha Reliability Coefficient for the TCPFS scale*

Factors	Cronbach's Alpha
Affective Attitude (AAT)	.90
Instrumental Attitude (IAT)	.93
Subjective Norm (SN)	.83
Controllability (CON)	.92
Self-Efficacy (SEF)	.86
Intention (INT)	.93
Behavior (BEH)	.86
TCPF Scale	.95

Predictive Model

Since the psychometric properties of the scale in the TPB model were acceptable, predictive (hypothesized) model was estimated by using structural equation modeling (Figure 3). Predictive model in the TPB framework helped to predict teachers' conceptions and behaviors regarding formative assessment in this study. The predictive (hypothesized) model demonstrated good fit (χ^2=2051.975, *df*=607, *p*<.001, RMSEA=.07, CFI=.88; TLI=.87, IFI= .88).The analysis showed acceptable statistical fit between the proposed model based on Theory of Planned Behavior and the observed data. The standardized regression weights of the paths from instrumental attitude, self- efficacy and controllability to intention were significant (*p*<.01). The strongest predictor of intention was self-efficacy (β= .55), followed by controllability (β= .16), and instrumental attitude (β= .16). However standardized regression weight of the path from subjective norm (β= -.04), and affective attitude (β= .07) to intention were not significant. The direct effect of intention and indirect effects of self-efficacy and controllability on behavior were also examined. In the predictive model, behavior (formative assessment practice) was predicted by both controllability (β= -.52)

and self-efficacy (β= .68). However, standardized regression weight of the path from intention to behavior was not significant (β= -.17). This model explained a substantial 71% of the variance in teachers' intentions and 15% of the variance in teachers' behaviors.

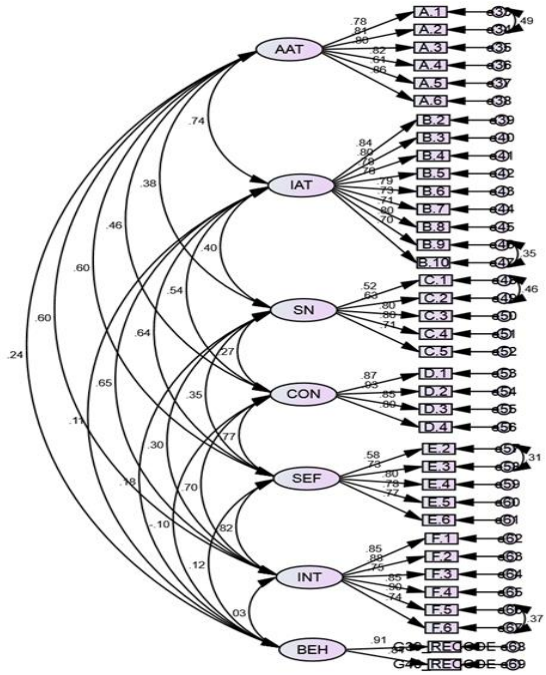


Figure 2. Confirmatory factor model based on the Theory of Planned Behavior

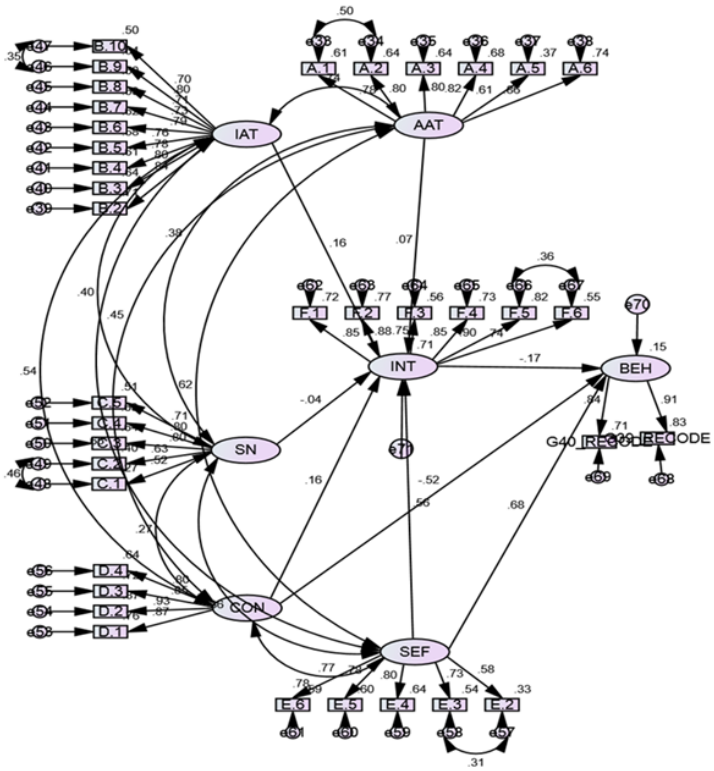


Figure 3. Predictive model of Theory of Planned Behavior

Discussion and Conclusion

The purpose of this study was twofold: 1) to adapt "Teachers' Conceptions and Practices of Formative Assessment Scale (TCPFS)" developed by Yan and Cheng (2015) to Turkish and examine psychometric features of the scale, 2) to apply the TPB model to examine teachers' conceptions and practices regarding formative assessment. First of all, linguistic validity of the scale was performed. The result showed that original version and Turkish version of the scale were quite similar in terms of linguistic fit. The translated instrument was applied to 400 primary school teachers for the construct validity of the scale. CFA was performed to examine whether the data support the proposed model of the scale. The fit index values of the structural model confirmed the construct validity of the scale. The reliability of the scale was examined with Cronbach's Alpha Coefficient. Reliability Coefficient indicated that the scale had good internal consistencies in the whole scale and also in the sub-scales. Therefore, validity and reliability of the adapted scale indicated that the scale has satisfactory psychometric features. In the current study, it was observed that some psychometric values for Turkish version of the scale were less than the values of original scale. These differences may occur while adapting educational and psychological instruments into different cultures and different languages (Hambleton, Merenda, & Spielberger, 2005; Geisinger, 1994; Sireci & Berberoğlu, 2000). In sum, adaptation and validation of the TCPFS scale to Turkish is important contribution to education research in Turkey. Given its parsimony and adequate fit, the adapted scale can be useful as a research and an assessment tool for measuring Turkish teachers' conceptions and practices of formative assessment.

In the second step of the study, the predictive model based on TPB was estimated. Structural equation modelling was used to examine the applicability of TPB in understanding teachers' conceptions and practices of formative assessment. The research results revealed that teachers' intentions to use formative assessment can be predicted by self-efficacy, perceived behavioral control, and instrumental attitude. Teachers' intentions were mostly predicted by self-efficacy. However, teachers' perceived behavioral control and instrumental attitude showed a weak impact on their intentions. This key finding was similar with research on TPB for teachers' assessment issues (Yan, 2014; Yan & Cheng, 2015). In the Yan and Cheng's (2015) study, teachers' self-efficacy had a stronger effect on their intentions than instrumental attitude to conduct formative assessment. Yan (2014) focused on school-based assessment issue by using TPB model. The researcher also found that self-efficacy had a higher impact on teachers' intentions to conduct school-based assessment than instrumental attitude. Studies demonstrated that teachers' self-efficacy can affect their classroom activities and student achievement (Gibson & Dembo, 1984), their teaching strategies (Allinder, 1994), and their attitudes toward new initiatives and reforms (Dixon & Haigh, 2009; Fuchs, Fuchs, & Bishop, 1992). The present study showed that teachers with higher level of self-efficacy tended to be more likely to conduct formative assessment. While self-efficacy, perceived behavioral control, and instrumental attitude were significant predictors of intention to conduct formative assessment, self-efficacy and perceived behavioral control were the significant predictors of formative assessment practices. Teachers' formative assessment behaviors were most strongly predicted by self-efficacy and strongly predicted by perceived behavioral control. The results showed that in general TPB model with self-efficacy has made great contribution to the prediction of intention and behavior. However, perceived behavioral control was less predictive than self-efficacy in the study. This finding was consistent with previous research on formative assessment conceptions and practices (Yan & Chen, 2015).

In the present study, TPB model explained higher percentage of variance (71%) in teachers' intentions to use formative assessment than in teachers' formative assessment practices (15%). Meta-analytic studies about TPB showed that the model worked well for predicting and explaining intentions and behaviors (Armitage & Conner, 2001; Notani, 1998; Sutton, 1998). Armitage and Conner (2001) analyzed 185 studies. They found that TPB model accounted, on average, 39% of the variance in intention and 27% of the variance in behavior. In Sutton (1998)'s meta-analytic review, TPB and Theory of Reasoned Action model (TRA) was compared. The author found that TPB model was greater than TRA model for predicting and explaining of intentions and behaviors. These models explained 40-50% of variance in intention and 19-38% of the variance in behavior. The current finding was consistent with the previous TPB studies. In conclusion, this study supported that TPB model is an appropriate theoretical model for understanding factors that may predict and explain teachers' intentions and practices regarding formative assessment.

However, the study also had some limitations. Most important limitation of the study was that the findings based on self-report data from the primary teachers regarding formative assessment practices. Self-report data may increase the possibility of participant bias. In future studies, not only self-report data but also observational data from teachers regarding formative assessment practices might be considered. Another limitation of the study was to use convenience sampling that might affect the generalization of the study.

References

Ajzen, I. (1985). *From intentions to actions: a theory of planned behavior*. In J. Kuhl, & J. Beckmann (Eds.), Action control: From cognition to behavior. Berlin: Springer Berlin Heidelberg.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50(2)*, 179-211.

Allinder, R. (1994). The relationship between efficacy and the instructional practices of special education teachers and consultants. *Teacher Education and Special Education, 17*(2), 86-95.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: a meta-analytic review. *British Journal of Social Psychology, 40* (4), 471-499.

Aydoğmuş, A. & Coşkun Keskin, S. (2012). Use of process based measurement and evaluation tools by social sciences teachers: İstanbul province sample. *Mersin University Faculty of Education Journal, 8 (2)*, 110-123.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman and Company.

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of crosscultural adaptation of self-report measures. *Spine, 25(24)*, 3186-3191.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18(1)*, 5-25.

Bentler, P.M. & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.

Black, P., & Wiliam, D. (2006). *Assessment for learning in the classroom*. In J. Gardner (Ed.), Assessment and learning, 9-25, SAGE Publication.

Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice, 11*, 301–318.

Brown, G.T.L., Harris, L.R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education 28*, 968-978.

Brown, G.T.L., & Gao, L. (2015). Chinese teachers' conceptions of assessment for and of learning: Six competing and complementary purposes. *Cogent Education, 2:* 993836.

Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy and Practice*, *16(3)*, 347–363.

Bracken, B. A. & Barona, A. (1991) State of the art procedures for translating, validating and using psychoeducational tests in crosscultural assessment. *School Psychology International*, 12 (1–2), 119–132.

Brislin, R. W. (1970) Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1 (3), 185–216.

Byrne B.M. (2001). *Structural equation modeling with AMOS. Basic concepts, applications, and programming*. Mahwah, N.J: Lawrence Erlbaum Associates.

Carles, D. (2015). Prospects for the implementation of assessment for learning. *Assessment in Education, 12(1),* 39–54.

Crawley, F. E. (1990). Intentions of science teachers to use investigative teaching methods: A test of the theory of planned behavior. *Journal of Research in Science Teaching. 27*(7), 623-716.

Conatser, P., Block, M., & Gansneder, B. (2002). *Aquatic instructors' beliefs toward* inclusion: The theory of planned behavior. *Adapted Physical Activity Quarterly, 19,* 172–187.

Conner, M., Warren, R., Close, S., & Sparks, P. (1999). Alcohol consumption and theory of planned behavior: An examination of the cognitive mediation of past behavior. *Journal of Applied Social Psychology, 29*,1676-1704.

Darling-Hammond, L., & Pecheone, R. (2010). *Developing an internationally comparable balanced assessment system that supports high-quality learning*. Princeton, NJ: Educational

Davis, D.S. & Neitzel, C. (2011). A self-regulated learning perspective on middle grades classroom assessment, *The Journal of Educational Research, 104(3),* 202-215.

Dixon, H., & Haigh, M. (2009). Changing mathematics teachers' conceptions of assessment and feedback. *Teacher Development, 13*(2), 173-186.

Ektem, I.S, & Keçici, S.E., & Pilten, G. (2016). Sınıf Öğretmenlerinin Süreç Odaklı Ölçme ve Değerlendirme Yöntemlerine İlişkin Görüşleri, *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi, 17(3)*, 661-680.

Frohbeiter, G., Greenwald, E., Stecher, B., & Schwartz, H. (2011). *Knowing and doing: What teachers learn from formative assessment and how they use information*. CRESST Report 802. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing.

Fullan, M. (1999). *Change forces: the sequel* (London, Falmer Press).

Fuchs, L., Fuchs, D., & Bishop, N. (1992). Instructional adaptations for students at risk. *Journal of Educational Research, 86*, 70-84.

Geisinger, K.F. (1994). Cross-cultural normative assessment: Translation and adaptation issues influencing the normative interpretation of assessment instruments. *Psychological Assessment, 6(4),* 304-312.

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76(2)*, 569-582.

Gong, B. (2010). *Using balanced assessment systems to improve student learning and school capacity: an introduction.* Council of Chief State School Officers and Renaissance Learning.

Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and*

psychological tests for cross-cultural assessment (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum.

Hambleton, R.K., Merenda, P.F., & Spielberger, C.D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum.

Haney, J. J., & McArthur, J. (2002). Four case studies of prospective science teachers' beliefs concerning constructivist teaching practices. *Science Education, 86*, 783–802.

Harrison, C. (2013). Collaborative action research as a tool for generating formative feedback on teachers' classroom assessment practice: the KREST project. *Teachers and Teaching: Theory and Practice, 19(2)*, 202-213.

Hattie, J. (2009). *Visible learning: a synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.

Heritage, M. (2007). Formative Assessment: What Do Teachers Need to Know and Do? *Phi Delta. Kappan, 89(2)*, 140–145.

Heritage, M. (2010a). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, CA: Corwin Press.

Heritage, M. (2010b). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Paper prepared for the Council of Chief State School Officers. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing (CRESST)

Herdman M, Fox-Rushby J, Badia X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Qual Life Res, 7*,323-335.

Huchting, K., Lac, A., & LaBrie, J.W. (2008). An application of the theory of planned behavior to sorority alcohol consumption. *Addictive Behavior, 33*, 538–551.

Jonsson, A., Lundahl, C. & Holmgren, A. (2015). Evaluating a large-scale implementation of Assessment for Learning in Sweden, *Assessment in Education: Principles, Policy & Practice, 22:1*, 104-121.

Kersaint, G., Lewis, J., Potter, R., & Meisels, G. (2007). Why teachers leave: Factors that influence retention and resignation. *Teaching and Teacher Education 23*, 775–794.

Konur, H., & Konur, B. (2011). İlköğretim öğretmenlerinin kullandıkları ölçme değerlendirme metotlarına ilişkin görüşleri. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (NEF-EFMED)*, *5*(2), 138-155.

Knight, P. (2002). A systemic approach to professional development: learning as practice. *Teaching and Teacher Education,* 18(3), 229–241.

Konur, H., & Konur, B. (2011). İlköğretim öğretmenlerinin kullandıkları ölçme değerlendirme metotlarına ilişkin görüşleri. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi (NEF-EFMED), 5(2), 138-155.*

Landis, R., Edwards, B. D., & Cortina, J. (2009).“Correlated residuals among items in the estimation of measurement models”. In C. E. Lance & R. J. Vandenberg (Eds.). Statistical and methodological myths and urban legends: Doctrine, verity, and fable in the organizational and social sciences (pp. 195-214). New York: Routledge.

MacCallum, R.C., Browne, M.W., Sugawara, H.M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. *Psychological Methods, 1*:130–49.

MoNE, (2013). *İlköğretim Kurumları Fen Bilimleri Dersi (3, 4, 5, 6, 7 ve 8. Sınıflar) Öğretim Programı*, Ankara: Devlet Kitapları Müdürlüğü.

National Council of Teachers of English. (2013). Formative assessment that truly informs instruction. Urbana, IL: Author. Retrieved from http://www.ncte.org/library/NCTEFiles/Resources/Positions/formativeassessment single.pdf

Notani, A. S. (1998). Moderators of perceived behavioral control's predictiveness in the theory of planned behavior: A meta-analysis. *Journal of Consumer Psychology*, *7*, 247-271.

Martin, J.J., & Kulinna, P.H. (2004). Self-Efficacy Theory and the Theory of Planned Behavior: Teaching Physically Active Physical Education Classes, *Research Quarterly for Exercise and Sport, 75:3,* 288-297

Martin, J.J., & Kulinna, P.H. (2005). A social cognitive perspective of physical activity related behavior. *Journal of Teaching in Physical Education, 24(3),* 265–281.

Öz., H. (2014). Turkish Teachers'Practices of Assessment for Learning in the English as a Foreign Language Classroom. *Journal of Language Teaching and Research, 5(4)*, 775-785.

Pajares, M. F. (1992). Teachers' beliefs and educational research: cleaning up a messy construct. *Review of Educational Research, 62*, 307-332.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44(1*), 57–84

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice, 5(1)*, 77–85.

Schifter, D. E. & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behavior. *Journal of Personality and Social Psychology*, *49*(3), 843-851.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shermis & Di Vesta (2011). *Classroom assessment in action*. Lanham, MD: Rowman & Littlefield Publishers.

Sireci, S. G. & Berberoğlu, G. (2000). Using bilingual respondents to evaluate translated-adapted items. *Applied Measurement in Education, 13*(3), 229-248.

Stanec, A. D. S. (2009). The theory of planned behavior: Predicting teachers' intentions and behavior during fitness testing. *Journal of Teaching in Physical Education, 28,* 255-271.

Stern, T. (2009). The Role of Formative Assessment in Enhancing Independent Learning and Reflective Teaching: Some Results of the Austrian IMST-Project. In K. Subramaniam & A. Mazumdar (Eds.), Proceedings of the epiSTEME-3-Conference (pp. 209-214). Mumbai: HBCSE Tata Institute of Fundamental Research.

Sumantri, M. S. & Satriani, R. (2016). The effect of formative testing and self-directed learning on mathematics learning outcomes. *International Electronic Journal of Elementary Education, 8 (3),* 507-524.

Sutton, S. (1998). Explaining and predicting intentions and behavior: How well are we doing? *Journal of Applied Social Psychology, 28*, 1318–1339.

Sümer, N. (2000). Yapısal eşitlik modelleri. *Türk Psikoloji Yazıları*, *3* (6), 49-74.

Şimşek, Ö. F. (2007). *Yapısal eşitlik modellemesine giriş: Temel ilkeler ve LISREL uygulamaları*. Ankara: Ekinoks Yayınevi.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson Education.

Deborah J. Terry, D.J., & O'Leary, J.E. (1995). The theory of planned behaviour: The effects of perceived behavioural control and self-efficacy. *British Journal of Social Psychology, 34*, 199-220.

Trumbull, E., & Lash, A. (2013). *Understanding Formative Assessment: Insights from Learning Theory and Measurement Theory*. San Francisco: WestEd

Volante, L., & Beckett, D. (2011). Formative assessment and the contemporary classroom: Synergies and tensions between research and practice. *Canadian Journal of Education, 34*, 239–255.

Yan, Z. (2014). Predicting teachers' intentions to implement school-based assessment using the Theory of Planned Behaviour. *Educational Research and Evaluation, 20*, 83-97.

Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education, 45*, 128-136.

Yau, K. (2004). *Teachers' perceptions of assessment for learning* (Unpublished master's thesis). University of Hong Kong, Pokfulam.

Yim, Y. S., & Cho, H.Y. (2016). Predicting pre-service teachers' intention of implementing peer assessment for low-achieving students. *Asia Pacific Educ. Rev., 17*:63–72.

Young, J.E.C. & Jackman, M.G.A. (2014). Formative assessment in the Grenadian lower secondary school: teachers'perceptions, attitudes and practices, *Assessment in Education: Principles, Policy & Practice, 21:4*, 398-411.

Zint, M. (2002). Comparing three attitude-behavior theories for predicting science teachers' intentions. *Journal of Research in Science Teaching, 39*(9), 819-844.

Wilson, M., & Draney, K. (2004). Some links between large-scale and classroom assessments: The case of the BEAR Assessment System. In M. Wilson (Ed.), *Toward coherence between classroom assessment and accountability.* The 103rd Yearbook of the National Society for the Study of Education, Part II (pp. 132–154). University of Chicago Press.

Woolfolk Hoy, A., Davis, H., & Pape, S. J. (2006). *Teacher knowledge and beliefs*. In P. A. Alexander, & P. H. Winne (Eds.), Handbook of educational psychology.

Appendix

Teachers' Conceptions and Practices of Formative Assessment Scale in Turkish Version

	A: DUYUŞSAL TUTUM ÖLÇEĞİ
1	Biçimlendirici değerlendirme yapmayı seviyorum.
2	Biçimlendirici değerlendirme zevkli bir süreçtir.
3	Biçimlendirici değerlendirme ilgi çeken bir değerlendirmedir.
4	Biçimlendirici değerlendirme öğretimi daha kolay hale getirmektedir.
5	Biçimlendirici değerlendirme öğrencilerin birbirlerine yardım etmelerini destekler.
6	Biçimlendirici değerlendirme daha iyi bir öğrenme ortamı olmasına yardımcı olur.
7	Biçimlendirici değerlendirme, bu sürece harcayacağım zamana değeceğini düşündüğüm bir değerlendirmedir.
	B: TUTUM ÖLÇEĞİ
8	Biçimlendirici değerlendirme öğrencilerin öğrenmeye yönelik merak duygularını arttırabilmektedir.
9	Biçimlendirici değerlendirme, öğrenci performanslarının doğru bir şekilde değerlendirmesini sağlayabilmektedir.
10	Biçimlendirici değerlendirme, öğrenme ve öğretimin birleşmesine yardımcı olabilmektedir.
11	Biçimlendirici değerlendirme, öğrencileri daha çok çalışmaya teşvik edebilmektedir.
12	Biçimlendirici değerlendirme öğrencilerin adil bir şekilde değerlendirilmelerine katkıda bulunabilmektedir.
13	Biçimlendirici değerlendirme ile öğrencilerin öğretmenlerinden aldıkları dönütler, öğrencilerin güçlü ve zayıf yönlerini öğrenmelerine yardımcı olabilmektedir.
14	Biçimlendirici değerlendirme öğrencilerin bağımsız öğrenmelerine yardımcı olabilmektedir.
15	Biçimlendirici değerlendirme, öğrencinin öğrenmede kendine olan güvenlerinin artmasına yardımcı olabilmektedir.
16	Biçimlendirici değerlendirme öğrenme-öğretme sürecinin niteliğini arttırabilmektedir.
17	Biçimlendirici değerlendirme öğretimde verimliliği arttırabilmektedir.
	C: ÖZNEL NORM ÖLÇEĞİ
Bildiğim kadarıyla, aşağıda belirtilen paydaşlar biçimlendirici değerlendirmenin kullanılmasının önemli olduğunu düşünürler.	
18	Milli Eğitim Bakanlığı'na bağlı çalışan memurlar.
19	Okulumun müdürü.
20	Öğrencilerimin velileri.
21	Öğrencilerim.
22	Meslektaşlarım.
	D: KONTROL EDİLEBİLİRLİK ÖLÇEĞİ
23	Biçimlendirici değerlendirmeyi ne kadar sıklıkta uygulayacağıma ben kendim karar verebilirim.
24	Biçimlendirici değerlendirmeyi ne zaman uygulayacağıma ben kendim karar verebilirim.
25	Biçimlendirici değerlendirmeyi uygulayıp uygulamayacağıma ben kendim karar verebilirim.
26	Biçimlendirici değerlendirmede uygulayacağım yönteme ben kendim karar verebilirim.
	E: ÖZ-YETERLİK ÖLÇEĞİ
27	Biçimlendirici değerlendirmeyi öğrenme-öğretme süreciyle birleştirebilirim.
28	Biçimlendirici değerlendirmeyi uygulamaya yönelik yeterli eğitimi aldım.

29	Biçimlendirici değerlendirme etkinlikleri oluşturabilirim.
30	Biçimlendirici değerlendirmeyi uygulamaya yönelik gerekli zamanı oluşturabilirim.
31	Biçimlendirici değerlendirmeyi uygulamaya yardımcı materyallere (CD, DVD, el kitabı gibi) sahibim.
32	Biçimlendirici değerlendirmeyi uygulayabilmek için gerekli becerilere sahibim.
	F: NİYET ÖLÇEĞİ
33	Biçimlendirici değerlendirme uygulamalarını denemek istiyorum.
34	Biçimlendirici değerlendirmeyi öğretimle birleştirmek istiyorum.
35	Uygun biçimlendirici değerlendirme etkinlikleri oluşturmak istiyorum.
36	Biçimlendirici değerlendirmeye yönelik ölçme ve değerlendirme yöntemlerini belirlemek istiyorum.
37	Biçimlendirici değerlendirmeyi uygulama konusunda gerekli çabayı göstereceğimi düşünüyorum.
38	Öğrencilerimi biçimlendirici değerlendirmeye katılımları konusunda cesaretlendirmek istiyorum.
	G: DAVRANIŞ ÖLÇEĞİ
39	Biçimlendirici değerlendirmeyi son altı ay içinde uyguladınız mı?
	1 ( ) Her gün 2 ( ) Neredeyse hergün 3 ( ) Çoğu gün 4 ( ) Bazı günler 5 ( ) Hiçbir zaman
40	Biçimlendirici değerlendirmeyi öğretim sürecinde son altı ayda ne kadar sıklıkta kullandınız?
	1 ( ) Çok sık 2 ( ) Sıklıkla 3 ( ) Çoğu gün 4 ( ) Bazı günler 5 ( ) Hiçbir zaman

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effect of Game-Assisted Mathematics Education on Academic Achievement in Turkey: A Meta-Analysis Study

Sedat Turgut [a,*], Özlem Dogan Temur [b]

Received: 14 August 2017
Revised: 20 October 2017
Accepted: 2 November 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236115

Abstract

In this research, the effects of using game in mathematics teaching process on academic achievement in Turkey were examined by meta-analysis method. For this purpose, the average effect size value and the average effect size values of the moderator variables (education level, the field of education, game type, implementation period and sample size) were calculated. MetaWin and Comprehensive Meta-Analysis (CMA) statistical programs were used for the analysis. Based on the inclusion criteria, 31 effect size values for 26 studies were calculated. Hedges's g coefficient was used when the effect sizes were calculated and the confidence level was accepted as 95%. The average effect size value was 0.792 with 0.077 standard error which was calculated by random-effects model. As a result, the effects of using game on academic achievement is medium and positive in mathematics teaching process.

Keywords: Game, mathematics, academic achievement, meta-analysis.

Introduction

It is the game itself which facilitates the child's interaction with the environment, completely opens the channels related to the communication. The teachers' use of games affects the perspectives of the children towards school and mathematics. The game has an important place in children's thinking on the numbers and initiating and maintaining mathematical communication (Trawick-Smith, Swaminathan, & Liu, 2016). The game improves oral communication, top-level social interaction skills, creative thinking skills, imaginary and divergent thinking skills and problem solving skills of the children (Wood & Attfield, 2005). It can be said that the game presents an environment to the children in which the communication process is practiced densely. From the social point of view, children's speaking, getting feedback to their questions in this process, communicating with their friends and teachers have importance in mathematics learning. The children comprehend the mathematical concepts before they use because mathematical thinking develops before language. Therefore, the proper use of mathematical words can help children to acquire the mathematical concepts. Using a clear and explanatory language during acquisition of mathematical knowledge and skills process of children is crucial (Presser, Clements, Ginsburg, & Ertle, 2015). The game can increase the effectiveness of teaching by generating a collaborative learning environment and creating discussion platforms. It also helps the students having less knowledge to improve their understanding (Ke, 2008).

In game process, the communication set by the children might increase their consciousness about mathematics. The more the variety of the scenarios and situations in educational environments are extended, the more the children gain consciousness about not only about their own but also the other children's mathematics. As long as the game based approaches are used in classrooms, the mathematical consciousness of children is expected to increase (Marcus, Perry, Dockett, & MacDonald, 2016). The students might not be relaxed if they perceive mathematics course as difficult. The game can change the students' perceptions that the mathematics is difficult and contribute them to feel relaxed in the course. The children can improve the informal mathematical knowledge they have acquired in game activities if they attend problem solving process (Brandth, 2013). It can be said that there is a relationship between the children's creating new structures with various materials during the game and their cognitive development. For instance, Wolfgang, Stannard, and Jones (2003) have stated that playing legos and making constructions with them improve cognitive development of children, moreover contribute learning the subjects requiring abstract thinking such as geometry, arithmetic, trigonometry in mathematics learning process.

With the development of the technology, the computer games have been involved in children's game world and the educators started using technology and technology-assisted games in learning process. Real life situations and experiences are learned in the best way in technology education. The students are able to both learn and maintain their learnings by practicing; the teachers are able to teach complex concepts more easily in technological setting (Bellamy & Mativo, 2010). Mind games are also used in learning-instruction process like computer games. Kazemi, Yektayar, and Abad (2012) stated that teaching chess improves the mathematical

[a,*] Corresponding author: Sedat Turgut, Bartin University, Faculty of Education, Bartin, Turkey. E-mail: sdtturgut42@hotmail.com
[b] Dumlupinar University, Faculty of Education, Turkey. E-mail: ozlemdtemur@gmail.com

problem solving abilities of students at different educational level considerably and contribute increasing the students' meta-cognitive abilities. The problem solving ability is a complex interaction between cognition and meta-cognition. The main source of troubles about problem solving might be that the students cannot follow their cognitive activities actively, cannot control them or do not have opportunity to supervise (Artzt & Armour-Thomas, 1992). The children learn new concepts, corroborate them by practicing, strengthen their mathematical skills and develop their problem solving strategies when they play games during the process of learning mathematics (Ernest, 1986). The teachers' use of games and creative pedagogical applications might be helpful for enhancing students' attitudes towards mathematics (Afari, Aldridge, Fraser, & Khine, 2013).

Reviewing the literature, there has not been found any researches examining the effect of using games in mathematics education by meta-analysis method. Therefore, it is thought that this research will provide a holistic perspective towards how teaching mathematics with games affect academic achievement. Thus, the current situation can be interpreted and some suggestions can be presented to future researches. It is aimed to statistically reveal the effect of using games in the process of teaching mathematics on academic achievement of the students. Additionally, it is tried to be determined whether the academic achievement gained as a consequence of using games in mathematics teaching process differs in terms of educational level, learning domain, type of the game, implementation period and sample size.

Methodology

Research Model

The effect of using games in the process of teaching mathematics on academic achievement has been examined through combining the findings obtained from primary studies. Meta-analysis technique has been used for this purpose. Meta-analysis is the statistical analysis of the quantitative data obtained from a number of independent studies on a specific subject and a method of overall evaluation about these studies (Glass, 1976; Lipsey & Wilson, 2001). The average effect sizes related to primary studies are identified, the relationship among these effect sizes and the relationship among study characteristics are evaluated with meta-analysis (Card, 2012). The effect size refers the degree of the relationship between two related variables, the size of the score emerging or to be found between the groups in an experimental implementation (Ellis, 2010). Meta-analysis includes standardization of various effect size statistics used in order to code different types of quantitative studies (Lipsey & Wilson, 2001). Thus, the numerical findings obtained from primary studies can be interpreted in a statistically coherent way. Moreover, various statistical errors of primary studies can be demonstrated. It can be commented that meta-analysis is a secondary analysis format.

As the findings about a certain research subject are interpreted by being combined in meta-analysis, it can be stated that meta-analysis also provides a basis to theory development (Hunter & Schmidt, 2004). Therefore, it can be determined what kind of researches are needed by making overall evaluations about the existing studies associated with the subject as well.

A series of phases that should be followed exists in meta-analysis studies. First of all, the problem is determined. Then the literature is scanned in accordance with the problem. As a consequence of this, the attained studies are coded in terms of the determined criteria. The statistical analysis of the data is done after this phase. Finally, the findings obtained as a result of the analysis are interpreted (Pigott, 2012; Sánchez-Meca & Marín-Martínez, 2010). The effect of using games in teaching mathematics on academic achievement has been examined within the frame of mentioned phases in this research.

Data Collection

The research data were collected in April 2017. The studies examining the effect of using games in teaching mathematics on academic achievement in Turkey established the data resources. YOK (Higher Education Council), ULAKBIM (National Academic Network and Information Center), Google Scholar databases were used so as to access the studies. For this purpose, the mentioned databases were scanned with the keywords "game and mathematics, mathematics and game, teaching mathematics with game". As a result, 60 works were reached. The studies to be included in the meta-analysis were determined according to the following criteria:

1. Studies should be prepared between the years of 2000-2017.

2. Studies should be written either in Turkish or in English.

3. Studies should be open to access in YOK, ULAKBIM and Google Academic databases.

4. Studies should be related to preschool, primary school, middle school, high school and university students who are studying in Turkey.

5. The studies should be experimental and the pretest-posttest control group model should be used in the studies.

6. In the studies, the experimental group should be taught with game(s) and the control group should be taught based on the traditional methods.

7. In the studies, statistical values such as sample sizes, arithmetic means, standard deviation etc. of both the experimental and the control groups should be given to calculate effect size.

30 studies were selected in accordance with these criteria. It has been noticed that 4 of the studies were the articles generated from dissertations or thesis and the articles were included in the study while 4 dissertations or thesis owning an article were ignored. Consequently, totally 26 studies, 4 articles and 22 dissertations, were included in the meta-analysis. As two different achievement tests

were implemented to the same group in a master's thesis from these studies, 2 effect size values of this study were calculated. As three different control groups were found in one of the master's thesis, 3 effect size values of this study were calculated. As there were also 2 different experimental groups in a doctoral dissertation, 2 effect size values were calculated for this study. In order that there is not a confusion, the letters a, b, c were added next to the study year of this type of studies when the analysis results were stated. Eventually, 31 effect size values were calculated related to 26 studies included in the meta-analysis.

Coding the Data

A coding form was constituted in accordance with the inclusion criteria to use at the phases of meta-analysis study. This form includes information about study number, study name, author name, study type, study year, sample size, educational level of sample, learning domain, implementation period, and sample size, arithmetic mean, standard deviation of the experimental and control groups, and validity and reliability information of the utilized assessment instruments. The information about the studies, to be included in the analysis, were coded by the two researchers in order to determine correctly and transfer the data without any error. After the coding, the two researchers compared the forms. As a result of the comparison, there was not found any difference between two researchers' coding. Consequently, the analysis phases were started.

Data Analysis

There are two main statistical approaches, fixed effect model and random effect model, in meta-analysis process (Borenstein, Hedges, Higgins, & Rothstein, 2009; Hunter & Schmidt, 2004). The distribution of the effect sizes of the studies included in meta-analysis is considered when the model to be used in analysis process is determined. In order to determine whether the distributions are homogenous, Q statistics can be used. Zero hypothesis indicating that all the studies included in meta-analysis share a mutual effect size with Q statistics is tested using chi-square distribution (Borenstein, Hedges, Higgins, & Rothstein, 2009). If the value obtained as a result of Q statistics is lower than p significance value and the value of degree of freedom (df) corresponding in chi-square (X^2) table, homogenous distribution is provided (Borenstein et. al, 2009). Otherwise (if it is higher than the value corresponding in chi-square table), it can be said that the distribution is heterogonous. If the effect sizes of the studies included in meta-analysis show homogenous distribution according to Q statistics, fixed effect model and if they show heterogonous distribution random effect model is preferred (Ellis, 2010). Another statistics that can be used for determining the distribution is I^2. I^2 statistics provides a ratio independent from size effect scale and its interpretation is heuristic (Higgins, Thompson, Deeks, & Altman, 2003). I^2 explains the heterogeneity effect in the analysis (Petticrew & Roberts, 2006). If the value calculated as a result of I^2 statistics is higher than 25%, it indicates heterogeneity; if it is 50%, it indicates medium heterogeneity and if it is 75%, it indicates high-level heterogeneity (Cooper, Hedges, & Valentine, 2009; Higgins, Thompson, Deeks, & Altman, 2003). Random effect model was used depending on the results of Q statistics and I^2 statistics (see Table 3) in this study.

Comprehensive Meta-Analysis (CMA) statistical programs were used for MetaWin, forest plot, funnel plot, publication bias, effect sizes and sub-group analysis for determining whether the effect sizes of the studies included in the study show normal distribution.

Funnel plot and Rosenthal's fail-safe N statistics were examined in finding publication bias associated with the studies. The fact that the effect sizes of the studies distribute around overall effect size symmetrically in funnel plot indicates lack of publication bias (Borenstein et. al., 2009). But an asymmetrical distribution might also remark a real heterogeneity (Tang & Liu, 2000). Therefore, it might not mean an absolute publication bias if the distribution is not completely symmetrical. Rosenthal's fail-safe N (FSN) statistics refers the number of new studies that should be added to the analysis in order to zero the effect size reached as a result of meta-analysis (Borenstein et. al., 2009). If FSN value (N) in the analysis is relatively higher than the observed studies, the results are enough resistant to publication bias (Rosenthal, 1991). In addition this, Mullen, Muellerleile, and Bryant (2001) produced N/(5k+10) (k is the number of included studies in meta-analysis) considering Rosenthal's fail-safe N statistics. In case that the value to be reached is higher than 1, they stated that the results are enough resistant to publication bias.

Hedge's g coefficient was used to calculate the effect sizes of the studies. Reliability degree was accepted as 95% in the calculations related to effect sizes. The criteria was taken into consideration that if it is between 0-0.20, it is weak; between 0.21-0.50, it is small; between 0.51-1.00, it is medium; and if it is higher than 1, it is large effect (Cohen, Manion, & Morrison, 2007, p. 521) while the effect sizes were interpreted. Educational level, learning domain, type of game, implementation period and sample size were identified as moderators in the study. Analog ANOVA test was utilized in the analysis of moderators.

The descriptive statistics of the studies examining the effect of game in the process of mathematics teaching on academic achievement in Turkey is presented in Table 1.

Table 1. *The Descriptive Statistics of the Studies Examining the Effects of Game on Academic Achievement in Mathematics*

		Frequency	Percentage (%)
Study Type	Article	4	15.38%
	Master's Thesis	19	73.07%
	Doctoral Dissertation	3	11.53%
Study Year	2004	2	7.69%
	2006	1	3.84%
	2007	2	7.69%
	2008	2	7.69%
	2009	1	3.84%
	2010	2	7.69%
	2011	1	3.84%
	2012	3	11.53%
	2013	2	7.69%
	2014	1	3.84%
	2015	3	11.53%
	2016	6	23.07%
Educational Level	Preschool	2	7.69%
	Elementary School	10	38.46%
	Middle School	13	50%
	Higher School	1	3.84%
Learning Domain	Mathematics	16	61.53%
	Geometry	5	19.23%
	Mathematics and Geometry	5	19.23%
Game Type	Computer-Assisted	7	26.92%
	Musical Game	1	3.84%
	Pedagogical Game	18	69.23%
Sample Size	1 -20 persons	-	-
	21-40 persons	6	23.07%
	41-60 persons	15	57.69%
	61 or more persons	5	19.23%
Implementation Period	1 -5 hours	2	7.69%
	6-10 hours	5	19.23%
	11-15 hours	3	11.53%
	16-20 hours	4	15.38%
	21-25 hours	-	-
	26-30 hours	1	3.84%
	31 or more hours	2	7.69%
	Unidentified	9	34.61%
Total		**26**	**100**

Observing Table 1, it is seen that four of the studies included in meta-analysis are articles (15.38%), 19 are master's theses (73.07%), and three are doctoral dissertations (23.07%). The largest number of the studies were done in 2016 (6 studies, 23.07%). It is seen that middle schools (13 studies, 50%) and elementary schools (10 studies, 38.46%) were focused in the studies in terms of educational levels. Within the aspect of learning domain, 16 of the studies are related to mathematics (61.33%), five are related to geometry (19.23%), five are related to both mathematics and geometry (19.23%). seven of the games used in the studies are computer-assisted (26.92%), one is musical (3.84%) and 18 are various pedagogical games (64.23%). It is seen that the largest number of studies were done in the range of 41-60 persons (15 studies, 57.69%), no study was done in the range of 1-20 persons. Among the identified studies, it is seen that most of the studies were implemented for 6-10 hours period (5 studies, 19.23%), there are nine studies (34.61%) in which the implementation period is not mentioned as hours.

Findings

Findings about the Effect of Game on Academic Achievement in Mathematics

Normal distribution plot has been observed in order to determine the convenience of combining effect sizes of 26 studies by meta-analysis. Normal distribution of the effect sizes of the studies is presented in Figure 1.

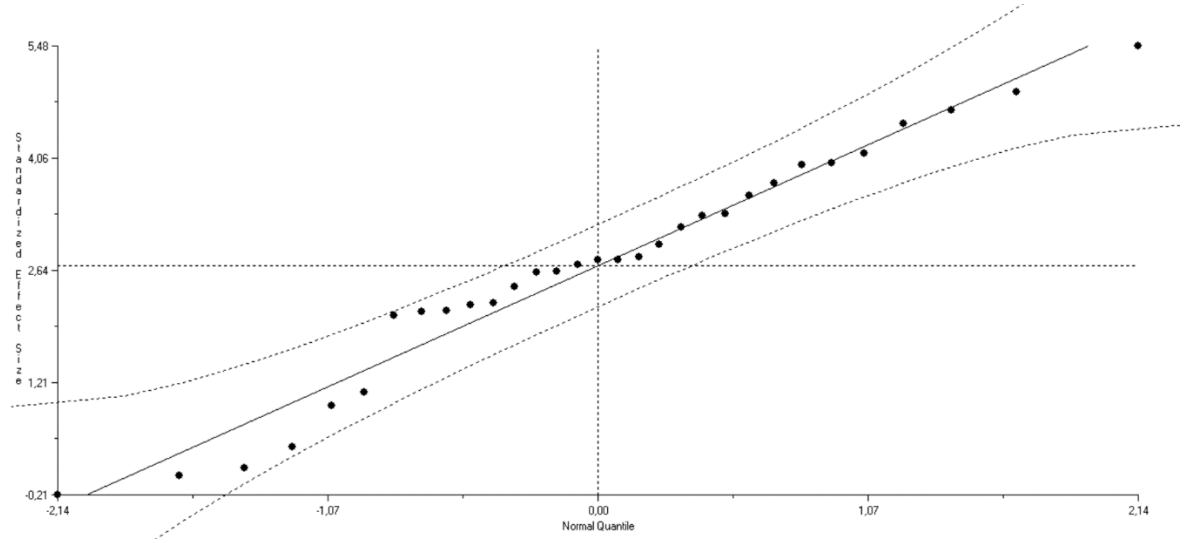


Figure 1. Normal distribution plot of the effect sizes of the studies included in meta-analysis

Looking at Figure 1, it is seen that the effect sizes of the studies distribute on the right and left sides of the normal distribution line and within the borders of confidence interval showed by dotted lines. Accordingly, it can be said that the effect sizes show normal distribution and they can be combined statistically by meta-analysis.

Before the effect sizes were calculated for the purpose of determining the effect of teaching by using games on academic achievement in mathematics, funnel plot of publication bias probability related to the studies included in meta-analysis was created and it is given in Figure 2.

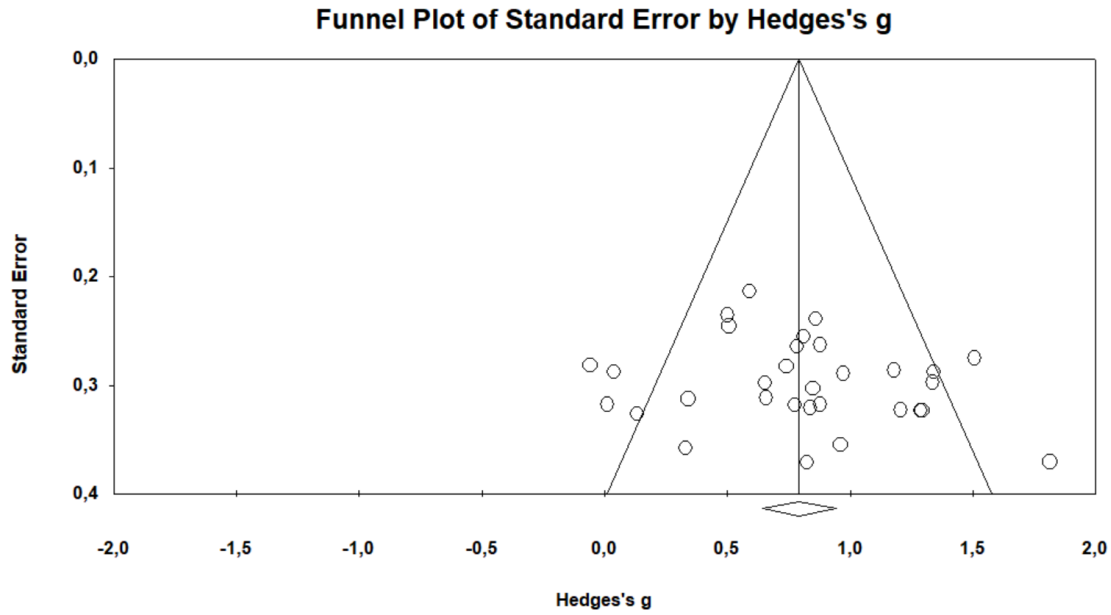


Figure 2. Funnel plot related to the effect sizes of the studies included in meta-analysis

Looking at Figure 2, it is seen that the effect sizes generally show an almost symmetrical distribution at the middle part of the funnel plot and at the right and left sizes of the line indicating combined effect size. In case of publication bias, it would be asymmetrical. In addition to this, the fact that the distribution is not completely symmetrical indicates publication bias. Therefore, Rosenthal's fail-safe N (FSN) was examined in addition to funnel plot. The related statistical information is given in Table 2.

Table 2. *Calculated Rosenthal FSN Statistics Results Related to Meta-analysis Examining the Effect of Using Game on Academic Achievement in Mathematics*

Bias Condition	
Z value for observed studies	15.02209
P value for observed studies	0.00000
Alpha	0.05
Direction	2
Z value for Alpha	1.95996
Number of Observed Studies	31
FSN	1791

Investigating Table 2, it is seen that N (FSN) value was calculated as 1791. According to N/(5k+10) formula suggested by Mullen, Muellerleile, and Bryant (2001), the result 1791/(5*31+10) is 10.8545. In terms of this result, it can be stated that the studies included in meta-analysis are resistant to publication bias.

Calculating the effect sizes of the studies to be included in meta-analysis; homogeneity value, average effect sizes and confidence intervals according to fixed effect and random effect models are presented in Table 3 in order to determine the model to be selected.

Table 3. *Average Effect Sizes and Lower and Upper Values of Confidence Interval According to Effect Model*

Model	Average Effect Size Value (ES)	95% Confidence Interval for Effect Size		Standard Error (SE)	Homogeneity Value (Q)	Degree of Freedom	I^2	*p*
		Lower Bound	Upper Bound					
Fixed	0.781	0.679	0.884	0.052	64.579	30	53.545	0.000
Random	0.792	0.641	0.944	0.077				

Looking at Table 3, homogeneity value of the studies in meta-analysis is calculated as Q= 64.579 in terms of fixed effect model. Critical value of 30 degree of freedom is 43.773 at 95% confidence level in chi-square table. According to this result, it is seen that Q value (64.579) is higher than the critical value corresponding 30 degree of freedom in chi-square table (x^2= 43.773 for df= 30). Depending on these findings, it can be stated that the studies analysed through meta-analysis show heterogeneous distribution. Additionally, I^2 value with 53.545% indicates heterogeneity over medium. Therefore, random effect model was chosen when the average effect sizes of the studies analysed through meta-analysis were calculated. The average effect size value was calculated as 0.792 with 0.077 standard error according to random effect model. The fact that the calculated effect size is positive indicates that treatment effect is on behalf of experimental groups. 0.792 effect size value reflects medium effect according to Cohen et. al. (2007). Depending on this reference, it can be inferred that using games in mathematics teaching affects academic achievement positively and this effect is at medium level.

Forest plot demonstrating the distribution of effect size values of primary studies examined through meta-analysis according to random effect model is presented in Figure 3. The squares seen in Figure 3 reflect the effect sizes of primary studies, the areas of the squares reflect the weight of the effect size of the study it belongs within the overall effect size. The numerical values of these weights are demonstrated at the rightmost part of the figure. The lines appearing at two sides of the squares represent upper and lower bounds of these effect sizes within 95% confidence interval. The equilateral quadrangle found at the lowest part of the squares indicates overall effect size. Investigating the effect sizes calculated, it is observed that the lowest effect value is -0.059 and the highest effect value is 1.815. Only 1 effect size value is negative among 31 effect sizes. Consequently, using games in the process of mathematics teaching affected on behalf of experimental groups in 30 studies.

Findings about Effect Sizes in Terms of Educational Level

Calculated effect sizes of the effect of using game in teaching mathematics on academic achievement in terms of educational level are presented in Table 4.

Table 4. *Effect Size Differences in Terms of Educational Level*

Variable	Homogeneity Value Between Groups (Q_B)	*p*	*n*	Average Effect Size Value (ES)	95% Confidence Interval for Effect Size		Standard Error (SE)
					Lower Bound	Upper Bound	
Educational Level	7.830	0.050					
Preschool			3	1.162	0.780	1.544	0.195
Elementary School			11	0.868	0.556	1.181	0.159
Middle School			16	0.661	0.492	0.829	0.086
Higher School			1	1.207	0.574	1.840	0.323

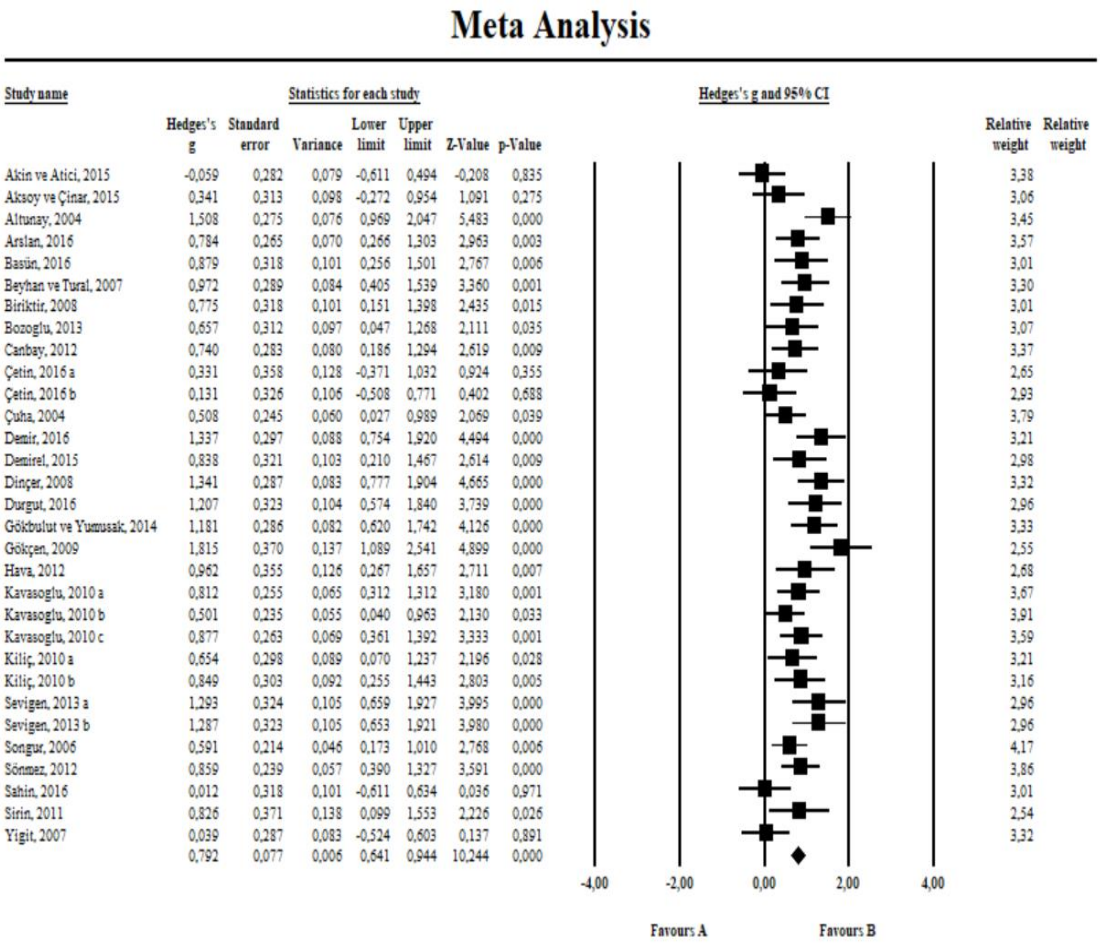


Figure 3. Forest plot of studies' effect sizes according to random effects model

Homogeneity value between groups in terms of educational level was calculated as (Q_B) 7.830. Critical value of 3 degree of freedom is 7.815 at 95% significance level in chi-square. It is seen that *Q* value is higher than the critical value corresponding to 3 degree of freedom in chi-square table (Q_B= 7.830, p= 0.050). Accordingly, a statistically significant difference between groups in terms of educational level is found. Regarding findings, using games in mathematics teaching has maximum effect on higher school and minimum effect on middle school in terms of educational level. According to Cohen et. al. (2007), the effect sizes calculated for higher school and preschool are large, the effect sizes calculated for primary school and middle school are medium-sized.

Findings about Effect Sizes in Terms of Learning Domain

Calculated effect sizes of the effect of using game in teaching mathematics on academic achievement in terms of learning domains are presented in Table 5.

Table 5. *Effect Size Differences in Terms of Learning Domain*

Variable	Homogeneity Value Between Groups (Q_B)	*p*	*n*	Average Effect Size Value (ES)	95% Confidence Interval for Effect Size		Standard Error (SE)
					Lower Bound	Upper Bound	
Learning Domain	0.618	0.734					
Mathematics			20	0.787	0.595	0.979	0.098
Geometry			5	0.905	0.591	1.219	0.160
Mathematics and Geometry			6	0.712	0.299	1.125	0.211

Homogeneity value between groups in terms of learning domain was calculated as (Q_B) 0.6188. Critical value of 2 degree of freedom is 5.991 at 95% significance level in chi-square table. It is seen that *Q* value is lower than the critical value corresponding to 2 degree of freedom in chi-square table (Q_B= 0.618, *p*= 0.734). Accordingly, a statistically significant difference between groups in terms of learrning domain is not found. Therefore, the academic achievement attained by using games does not show a statistically significant differentiation in terms of different learning domains. According to Cohen et. al. (2007), the effect sizes calculated for learning domains have medium effect.

Findings about the Effect Sizes in Terms of Game Types

Calculated effect sizes of the effect of using game in teaching mathematics on academic achievement in terms of game types are presented in Table 6.

Homogeneity value between groups in terms of game types was calculated as (Q_B) 6.667. Critical value of 2 degree of freedom is 5.991 at 95% significance level in chi-square.

Table 6. *Effect Size Differences in Terms of Game Types*

Variable	Homogeneity Value Between Groups (Q_B)	*p*	*n*	Average Effect Size Value (ES)	95% Confidence Interval for Effect Size		Standard Error (SE)
					Lower Bound	Upper Bound	
Game Type	6.667	0.036					
Computer-Assisted			7	0.472	0.091	0.854	0.195
Pedagogical Game			23	0.859	0.711	1.008	0.076
Musical Game			1	1.341	0.777	1.904	0.287

It is seen that *Q* value is higher than the critical value corresponding to 2 degree of freedom in chi-square table (Q_B= 6.667, *p*= 0.036). Accordingly, a statistically significant difference between groups in terms of game types is found. Regarding the findings, musical games have maximum effect on academic achievement in mathematics teaching. According to Cohen et. al. (2007), the effect sizes calculated for computer-assisted games are low, the effect sizes calculated for pedagogical games are medium, and the effect sizes calculated for musical games are large.

Findings about the Effect Sizes in Terms of Implementation Period

Calculated effect sizes of the effect of using game in teaching mathematics on academic achievement in terms of implementation period are presented in Table 7.

Table 7. *Effect Size Differences in Terms of Implementation Period*

Variable	Homogeneity Value Between Groups (Q_B)	*p*	*n*	Average Effect Size Value (ES)	95% Confidence Interval for Effect Size		Standard Error (SE)
					Lower Bound	Upper Bound	
Implementation Period	8.313	0.216					
1-5 hours			2	0.337	-0.269	0.942	0.309
6-10 hours			7	1.010	0.655	1.365	0.181
11-15 hours			3	0.797	0.460	1.134	0.172
16-20 hours			4	0.849	0.566	1.132	0.145
21-25 hours			-	-	-	-	-
26-30 hours			1	0.341	-0.272	0.954	0.313
31 or more hours			2	0.349	-0.211	0.910	0.286
Unidentified			12	0.837	0.560	1.115	0.142

Looking at Table 7, it is seen that there is not a study in 21-25 hours implementation period range whereas implementation period is not identified in 9 studies. Homogeneity value between groups in terms of implementation period was calculated as (Q_B) 8.313. Critical value of 6 degree of freedom is 12.592 at 95% significance level in chi-square table. It is seen that Q value is lower than the critical value corresponding to 6 degree of freedom in chi-square table (Q_B= 8.313, p= 0.216). Accordingly, a statistically significant difference between groups in terms of implementation period is not found. Therefore, the academic achievement attained by using games in different implementation periods does not show a statistically significant differentiation. According to Cohen et. al. (2007), the effect sizes calculated for implementation periods of 1-5 hours, 26-30 hours and 31 or more hours are low, implementation periods of 11-15 hours, 16-20 hours and unidentified ones are medium, and the effect sizes calculated for implementation periods of 6-10 hours are large.

Findings about the Effect Sizes in Terms of Sample Size

Calculated effect sizes of the effect of using game in teaching mathematics on academic achievement in terms of sample size are presented in Table 8.

Table 8. *Effect Size Differences in Terms of Sample Size*

Variable	Homogeneity Value Between Groups (Q_B)	p	n	Average Effect Size Value (ES)	95% Confidence Interval for Effect Size		Standard Error (SE)
					Lower Bound	Upper Bound	
Sample Size	1.016	0.602					
1-20 persons			-	-	-	-	-
21-40 persons			7	0.612	0.162	1.063	0.230
41-60 persons			17	0.861	0.661	1.061	0.102
61 or more persons			7	0.788	0.545	1.030	0.124

Investigating Table 8, it is seen that there is not a study in 1-20 persons sample size. Homogeneity value between groups in terms of implementation period was calculated as (Q_B) 1.016. Critical value of 2 degree of freedom is 5.991 at 95% significance level in chi-square table. It is seen that Q value is lower than the critical value corresponding to 2 degree of freedom in chi-square table (Q_B= 1.016, p= 0.602). Accordingly, a statistically significant difference between groups in terms of implementation period is not found. Therefore, the academic achievement attained by using games with different sample sizes does not show a statistically significant differentiation. According to Cohen et al. (2007), the effect sizes calculated for sample size are low.

Discussion, Conclusion and Recommendations

In this meta-analysis study examining the effect of using games in the process of teaching mathematics, 31 effect sizes of individual studies was calculated. 30 of these values are positive while 1 of them is negative. The average effect value calculated according to random effect model is 0.792. This value reflects a low effect according to Cohen et. al. (2007). It has been observed that 4 studies (Akın & Atıcı, 2015; Çetin, 2016b; Şahin, 2016; Yiğit, 2007) have low effect, 4 studies (Aksoy & Çınar, 2015, Çetin, 2016a; Çuha, 2004; Kavasoğlu, 2010b) have small effect, 15 studies (Arslan, 2016; Başün, 2016; Beyhan & Tural, 2007; Biriktir, 2008; Bozoğlu, 2013; Canbay, 2012; Demirel, 2015; Hava, 2012; Kavasoğlu, 2010a; Kavasoğlu, 2010c; Kılıç, 2010a; Kılıç, 2010b; Songur, 2006; Sönmez, 2012; Şirin, 2011) have medium level of effect, and 8 studies (Altunay, 2004; Demir, 2016; Dinçer, 2008; Durgut, 2016; Gökbulut & Yumuşak, 2014; Gökçen, 2009; Sevigen 2013a; Sevigen 2013b) have large effect according to the calculated effect size values related to primary studies analysed through meta-analysis. Accordingly, it can be inferred that using game in mathematics teaching process generally effects academic achievement positively. This result shows similarity with the literature. Ku, Chen, Wu, Lao, & Chan (2014) stated that game based learning is more effective than pencil-paper based (traditional) learning processes and the students feel more comfortable and their performances enhance in mathematics courses assisted by game based learnings. Uğurel (2003) remarked that mathematics teaching by using games and activities increases interest, provides motivation, offers the opportunity of active participation and permanent learning by providing use of all sense organs in learning process. Yılmaz (2014) expressed that teaching through games positively effects the students' academic achievement and their attitudes towards mathematics course. Özgenç (2010) stated that game based mathematics activities affect the interest of the students in the course and their participation positively, enhance their interaction with both their teachers and friends. Fırat (2011) reported that mathematics teaching maintained by computer-assisted educational games has a positive effect on students' learning the concept of probability.

Educational level, learning domain, game type, sample size and implementation period related to the primary studies were determined as moderators and the effect sizes in terms of these values were calculated. Thus, it was examined whether the academic achievement attained by using game in the process of teaching mathematics differs statistically according to moderators.

The effect sizes of four different groups including preschool, primary school, middle school, higher school were calculated in terms of educational level in the analysis. Consequently, it was found out that maximum effect of using game in mathematics teaching process on

academic achievement is at higher school while minimum effect is at middle school level. The fact that there is only one study at higher school level might be effective in existence of a result like this.

The effect sizes of three different groups including mathematics, geometry and mathematics and geometry were calculated in terms of learning domains. Consequently, it was found out that the effect of using game in mathematics teaching does not differentiate according to learning domains, has similar effects in all three groups.

The effect sizes of three different groups including computer-assisted, pedagogical and musical games were calculated in terms of game type. Consequently, it was found out that using musical games has the maximum effect on achievement. The fact that there is only one musical game among the studies done by meta-analysis might have effected this result. However, some studies supporting this finding are encountered in the literature. For instance, Yılmaz (2006) stated that musical game activities are effective on 5-6 aged students' acquisition of number and operation concepts in his study.

The effect sizes of seven groups including 1-5 hours, 6-10 hours, 11-15 hours, 16-20 hours, 26-30 hours, 31 or more hours and unidentified were calculated in terms of implication period. Consequently, it was observed that the effect of using game in mathematics teaching does not differentiate according to implementation period.

The effect sizes of three groups including 21-40 persons, 41-60 persons, 61 or more persons were calculated in terms of sample size. Consequently, it was observed that the effect of using game in mathematics teaching does not differentiate according to sample size.

The effect of using game in the process of mathematics education on only academic achievement was examined in terms of different variables in this study. The effect of using game on motivation, attitude, achievement motive, achievement permanency and acquisition of concepts can be investigated in further researches.

The effect size of only one study at higher school level was calculated in terms of educational level in the analysis. No effect size at high school or faculty level was calculated. The effect size of only one study including musical game was calculated in terms of game type, as well. Accordingly, it can be argued that mentioned types of studies are needed to entirely determine the effect of using game in mathematics teaching. Additionally, it can be pointed out that the statistical values needed for meta-analysis should be given completely both in existing and future studies.

References

The references marked with asterisk (*) demonstrate the studies included in meta-analysis.

Afari, E., Aldridge, J. M., Fraser, B. J., & Khine, M. S. (2013). Students' perceptions of the learning environment and attitudes in game-based mathematics classrooms. *Learning Environments Research*, *16*(1), 131-150.

*Akkuş-Sevigen, F. (2013). *Oyun Temelli Matematik Eğitim Programı'nın çocuğun matematik gelişimine etkisinin incelenmesi* (Unpublished master's thesis). Gazi University, Ankara.

*Aksoy, N. C. ve Çınar, C. (2015). Dijital oyun tabanlı matematik öğretiminin 6. sınıf öğrencilerinin başarı ve başarı güdüsüne etkisi. *Journal of Educational Sciences,* 167-177.*Altunay, D. (2004). *Oyunla desteklenmiş matematik öğretiminin öğrenci erişişine ve kalıcılığa etkisi* (Unpublished master's thesis). Gazi University, Ankara.

*Altunay, D. (2004). *Oyunla desteklenmiş matematik öğretiminin öğrenci erişişine ve kalıcılığa etkisi* (Unpublished master's thesis). Gazi University, Ankara.

*Arslan, N. (2016). *Oyun destekli öğretimin 5. sınıf temel geometrik kavramlar ve çizimler konusunun öğretiminde öğrencilerin başarısına etkisi* (Unpublished master's thesis). Sakarya University, Sakarya.

Artzt, A., & Armour-Thomas, E. (1992). Development of a cognitive-meta-cognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction*, *9*(2), 137-175.

*Aslan-Akın, F. ve Atıcı, B. (2015). Oyun tabanlı öğrenme ortamlarının öğrenci başarısına ve görüşlerine etkisi. *Turkish Journal of Educational Sciences, 2*(2), 75-102.

*Başün, A. R. (2016). *Oyunla öğretimin çarpanlar ve katlar alt öğrenme alanında başarı ve kalıcılığa etkisi* (Unpublished master's thesis). Ondokuz Mayıs University, Samsun.

Bellamy, J. S., & Mativo, J. M. (2010). A different angle for teaching math. *Technology Teacher*, *69*(7), 26-28.

*Beyhan, N. ve Tural, H. (2007). İlköğretim matematik öğretiminde oyunla öğretimin erişiye etkisi. *Dokuz Eylül Üniversitesi Buca Eğitim Fakültesi Dergisi, 21,* 37-48.

*Biriktir, A. (2008). *İlköğretim 5. sınıf matematik dersi geometri konularının verilmesinde oyun yönteminin erişiye etkisi* (Unpublished master's thesis). Selçuk University, Konya.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis.* UK: John Wiley & Sons, Ltd., Publications.

*Bozoğlu, U. (2013). *Ortaokul 7. sınıf matematik dersi alan-çevre ilişkisi konusunda oyun temelli öğretimin öğrenci başarısına etkisi* (Unpublished master's thesis). Ondokuz Mayıs University, Samsun.

Brandt, B. (2013). Everyday pedagogical practices in mathematical play situations in German "kindergarten". *Educational Studies in Mathematics, 84*(2), 227-248.

*Canbay, İ. (2012). *Matematikte eğitsel oyunların 7. sınıf öğrencilerinin öz-düzenleyici öğrenme stratejileri, motivasyonel inançları ve akademik başarılarına etkisinin incelenmesi* (Unpublished master's thesis). Marmara University, İstanbul.

Card, N. A. (2012). *Applied meta-analysis for social science researh.* London: The Guilford Press.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education.* New York: Routledge.

Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis.* New York: Russell Sage Publication.

*Çetin, Ö. (2016). *Ortaokul öğrencilerinin matematiksel oyun geliştirme süreçlerinin başarı, tutum ve problem çözme stratejilerine etkisi* (Unpublished doctoral dissertation). Necmettin Erbakan University, Konya.

*Çuha, S. S. (2004). *Matematik öğretiminde eğitsel oyunların başarı, akademik benlik, başarı güdüsü ve kalıcılık üzerindeki etkisi* (Unpublished master's thesis). Hacettepe University, Ankara.

*Demir, M. R. (2016). *Farklı oyun türlerine dayalı matematik öğretiminin 1. sınıf öğrencilerinin erişi ve kalıcılık düzeylerine etkisi* (Unpublished doctoral dissertation). Necmettin Erbakan University, Konya.

*Demirel, T. (2015). *Zekâ oyunlarının Türkçe ve matematik derslerinde kullanılmasının ortaokul öğrencileri üzerindeki bilişsel ve duyuşsal etkilerinin değerlendirilmesi* (Unpublished doctoral dissertation). Atatürk University, Erzurum.

*Dinçer, M. (2008). *İlköğretim okullarında müziklendirilmiş matematik oyunlarıyla yapılan öğretimin akademik başarı ve tutuma etkisi* (Unpublished master's thesis). Abant İzzet Baysal University, Bolu.

*Durgut, A. (2016). *Meslek yüksekokulu öğrencileri için eğitsel matematik oyunu geliştirilmesi ve başarıya etkisinin incelenmesi* (Unpublished master's thesis). Balıkesir University, Balıkesir.

Ellis, P. D. (2010). *The essential guide to effect sizes. Statistical power, meta-analysis, and the interpretation of research result.* New York: Cambridge University Press.

*Erkin-Kavasoğlu, B. (2010). *İlköğretim 6, 7 ve 8. Sınıf matematik dersinde olasılık konusunun oyuna dayalı öğretiminin öğrenci başarısına etkisi* (Unpublished master's thesis). Gazi University, Ankara.

Ernest, P. (1986). Games. A rationale for their use in the teaching of mathematics in school. *Mathematics in School*, *15*(1), 2-5.

Fırat, S. (2011). *Bilgisayar destekli eğitsel oyunlarla gerçekleştirilen matematik öğretiminin kavramsal öğrenmeye etkisi* (Unpublished master's thesis). Adıyaman University, Adıyaman.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of researh. *Educational Researcher, 5*(10), 3-8.

*Gökbulut, Y. ve Yücel-Yumuşak, E. (2014). Oyun destekli matematik öğretiminin 4. sınıf kesirler konusundaki erişi ve kalıcılığa etkisi. *Turkish Studies, 9*(2), 673-689.

*Gökçen, E. (2009). *Ortak bölenler ve katlar konusunun oyun ile öğretiminin başarıya etkisi* (Unpublished master's thesis). Çanakkale Onsekiz Mart University, Çanakkale.

*Hava, K. (2012). *Eğitsel bilgisayar oyunu tasarlama yönteminin, ilköğretim 4. sınıf öğrencilerinin akademik başarısına etkisi* (Unpublished master's thesis). Gazi University, Ankara.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327*(7414), 557-560.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis. Correcting error and bias in research findings.* Thousand Oaks, CA: Sage publications.

Kazemi, F., Yektayar, M., & Abad, A. M. B. (2012). Investigation the impact of chess play on developing meta-cognitive ability and math problem-solving power of students at different levels of education. *Procedia-Social and Behavioral Sciences*, *32*, 372-379.

Ke, F. (2008). Alternative goal structures for computer gamebased learning. *International Journal of Computer-Supported Collaborative Learning, 3*(4), 429-445.

Ku, O., Chen, S. Y., Wu, D. H., Lao, A. C. C., & Chan, T.-W. (2014). The effects of game-based learning on mathematical confidence and performance: High ability vs. low ability. *Journal of Educational Technology & Society*, *17*(3), 65-78.

Lipsey, M. W.. & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage publications.

Marcus, A., Perry, B., Dockett, S., & MacDonald, A. (2016). Children noticing their own and others' mathematics in play. In B. White, M. Chinnappan, & S. Trenholm (Eds.). *Opening up mathematics education research* (Proceedings of the 39th annual conference of the Mathematics Education Research Group of Australasia) (pp. 439-446). Adelaide: MERGA.

Mullen, B., Muellerleile, P., & Bryant, B. (2001). Cumulative meta-analysis: A consideration of indicators of sufficiency and stability. *Personality and Social Psychology Bulletin, 27*(11), 1450-1462.

Özgenç, N. (2010). *Oyun temelli matematik etkinlikleriyle yürütülen öğrenme ortamlarından yansımalar* (Unpublished master's thesis). Karadeniz Teknik University, Trabzon.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences. A practical guide.* Malden, MA: Blackwell Publishing.

Pigott, T. (2012). *Advances in meta-analysis*. New York: Springer Science & Business Media.

Presser, A. L., Clements, M., Ginsburg, H., & Ertle, B. (2015). Big math for little kids: The effectiveness of a preschool and kindergarten mathematics curriculum. *Early Education and Development, 26*(3), 399-426.

Rosenthal, R. (1991). *Meta-analytic procedures for social research.* California: Sage Publication.

Sánchez-Meca, J., & Marín-Martínez, F. (2010). Meta analysis. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 7, pp. 274-282). Oxford: Elsevier.

*Songur, A. (2006). *Harfli ifadeler ve denklemler konusunun oyun ve bulmacalarla öğrenilmesinin öğrencilerin matematik başarı düzeylerine etkisi* (Unpublished master's thesis). Marmara University, İstanbul.

*Şahin, H. B. (2016). *Eğitsel bilgisayar oyunlarıyla destekli matematik öğretiminin öğrencilerin akademik başarılarına ve duyuşsal özelliklerine etkisi* (Unpublished master's thesis). Eskişehir Osmangazi University, Eskişehir.

*Şirin, S. (2011). *Anaokuluna devam eden beş yaş grubu çocuklara sayı ve işlem kavramlarını kazandırmada oyun yönteminin etkisi* (Unpublished master's thesis). Uludağ University, Bursa.

Tang, J. L., & Liu, J. L. Y. (2000). Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology, 53*(5), 477-484.

Trawick-Smith, J., Swaminathan S., & Liu, X. (2016) The relationship of teacher-child play interactions to mathematics learning in preschool. *Early Child Development and Care*, *186*(5), 716-733.

*Tural-Sönmez, M. (2012). *6. sınıf matematik derslerinde web üzerinden sunulan eğitsel matematik oyunlarının öğrenci başarısına etkisi* (Unpublished master's thesis). Çukurova University, Adana.

Uğurel, I. (2003). *Ortaöğretimde oyunlar ve etkinlikler ile matematik öğretimine ilişkin öğretmen adayları ve öğretmenlerin görüşleri* (Unpublished master's thesis). Dokuz Eylül University, İzmir.

Wolfgang, C., Stannard, L., & Jones, I. (2003). Advanced constructional play with LEGOs among preschoolers as a predictor of later school achievement in mathematics. *Early Child Development and Care*, *173*(5), 467-475.

Wood, E., & Attfield, J. (2005). *Play, learning and the early childhood curriculum*. London: Paul Chapman Publishing.

Yılmaz, E. (2006). *Okul öncesi eğitim kurumlarına devam eden 6 yaş çocuklarının sayı ve işlem kavramlarını kazanmalarında oyun etkinliklerinin kullanılmasının etkisi* (Unpublished doctoral dissertation). Çukurova University, Adana.

Yılmaz, D. (2014). *Ortaokul 5. sınıf matematik dersi geometrik cisimler öğretiminde, matematik oyunları kullanımının öğrenci başarısı ve tutumuna etkisi* (Unpublished master's thesis). Gazi University, Ankara.

*Yiğit, A. (2007). *İlköğretim 2. sınıf seviyesinde bilgisayar destekli eğitici matematik oyunlarının başarıya ve kalıcılığa etkisi* (Unpublished master's thesis). Çukurova University, Adana.

*Zaif-Kılıç, A. (2010). *İlköğretim 1. sınıf matematik dersindeki işlem becerilerinin kazandırılmasında oyunla öğretimin başarıya etkisi* (Unpublished master's thesis). Celal Bayar University, Manisa.

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

Examining South Korea's Elementary Physical Education Performance Assessment Using Assessment Literacy Perspectives

Yongnam Park [a]

Received: 3 October 2017
Revised: 28 November 2017
Accepted: 29 November 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236116

Abstract

This study examines the issues pertaining to South Korea's elementary physical education (PE) performance assessment, using an assessment literacy (Hay & Penney, 2013) perspective to propose future directions. Eight elementary teachers currently teaching PE were selected as participants. Data were collected through semi-structured in-depth interviews and on-site data analysis, and analyzed based on an inductive categorical analysis, the results of which were elicited from the four concept factors of assessment literacy. Four themes presented themselves during the research: first, teachers were unclear about the concept of assessment, making it difficult for them to carry out assessments effectively; second, assessments were conducted in a labor exchange and recycling manner, reducing their effectiveness; third, there was a lack of feedback; and fourth, there were teachers' critical thinking without pedagogy. These are the main problems in assessing PE in elementary schools. As for future directions, this study proposes the need for diversifying current teachers' education geared towards enhancing their assessment literacy capability, for providing on-site guidance to build students' assessment literacy, and for evaluating the assessment procedure.

Keywords: Assessment, Assessment literacy, Physical education assessment, Performance assessment.

Introduction

As accountability in education is being increasingly emphasized, more importance has been attached to assessment, whether at the national, regional, or school level (Hardman & Marshall, 2000). South Korea has been highly ranked each year in the OECD-led Programme for International Student Assessment (PISA), and it has made assessments the focus of its education system to the extent that it is sometimes called "the Evaluation Republic." Subject evaluation activities are used to determine teaching quality; not only teachers, but also the students and their parents are evaluators.

In South Korea's evaluation-oriented school education, however, PE has been ignored and is in fact classified as optional; it is not considered for college entrance exams, and does not lend itself to performance evaluations, because of its unique characteristics. In this respect, standards for PE achievements and evaluations were promulgated, and research on developing an evaluation system was conducted at a national level, led by the Korea Institute for Curriculum and Evaluation (KICE). Performance assessment is an evaluation method that observes and judges students' knowledge, function and attitude in various authentic tasks and situations (McMillan, 2007). During the performance assessment process, teachers make comprehensive judgments through students' answers, outputs, and behaviors through several measurements and observations. Even with such national efforts, teachers still face difficulties with the evaluation of PE classes in their teaching activities (Yoo, 2005).

These difficulties have also been reported outside Korea. Issues such as teachers' indifference in assessment, their lack of knowledge about assessment and their attitude in failing to recognize assessment as part of PE have consistently been raised in various PE assessment-specific studies (Annerstedt & Larsson, 2010; Hay & Penney, 2013; Matanin & Tennehill, 1994). In particular, elementary school teachers assigned to teach various subjects are experiencing relatively more difficulties in PE assessment, and it is vital to galvanize research and interest in this area. Studies on the extent of teachers' knowledge regarding assessment and how well they can be implemented and interpreted also need to be conducted (Thompson & Penney, 2015).

Various levels of recommendations thus far made regarding PE evaluations have been fragmented, and do not provide clear indications about the knowledge, techniques, abilities, and attitudes that are necessary for appropriately evaluating PE. By positioning PE assessments into social and cultural activities instead of treating them as measurement-oriented assessments from a scientific perspective, the assessment literacy concept recently proposed by Hay and Penney (2013) has garnered attention; it provides an alternate perspective, incorporating the knowledge, techniques, and attitudes

[a] Correspondence details: Yongnam Park, Department of Physical Education, Seoul National University, KOREA 71, 1 Gwanak-ro, Seoul, 08826, KOREA. Phone: (+82) 10-2976-5845. E-mail: mri4558@snu.ac.kr

contained in the assessment.

Assessment literacy generally refers to an ability to understand and utilize the information obtained from developing and evaluating assessment standards. This is indispensable for appraising teachers' understanding of their assessment procedures and assignments, and the quality of students' performance (Fullan, 2002). In addition, the PE assessment literacy concept proposed by Hay and Penney (2013) emphasizes understanding assessment procedures and being aware of both the efforts involved in requesting assessments and responding to the assignments, and the non-educational intentions and latent outcomes implied in the assessment procedures. Assessment literacy comprises four concepts: assessment comprehension, assessment application, assessment interpretation, and critical engagement with assessment. Assessment comprehension is related to teachers' knowledge and understanding of the conditions for achievement standards or efficient evaluations; assessment application concerns conducting assessments (including both teacher evaluation and student evaluation); assessment interpretation considers the social roles and interrelations of evaluations; and critical engagement with assessment refers to the capability needed for a natural assessment plan, implementation, and results, while recognizing the outcomes and influences of the assessments. The ideas behind assessment literacy represent a holistic perspective as to what teachers are aware of and how they conduct, interpret, critique, and conceptualize assessments in sociocultural contexts; this helps foster a holistic assessment instead of an ad-hoc approach to individual issues. As seen in the analysis of the four conceptual factors of PE assessment literacy, they constitute a theoretical framework for elucidating teachers' expertise on assessments and the process of their evaluations; performance assessment can be conducive to exposing issues in PE assessments since they are a representative evaluation method. Therefore, this study examines the performance assessment issues in elementary school PE, with the intent of providing guidelines for future training for effective assessment.

Material & Methodology

Participants

The research participants were selected via the purposeful sampling method (Creswell, 2009). First, an initial pool was created with full-time teachers having five or more years of experience in elementary school teaching in South Gyeongsang Province, who were recommended by their colleagues. Second, the teachers from the initial pool who met certain specified criteria were selected as final research participants. These criteria were – the teachers should currently be teaching PE and conducting assessments as lead or full-time teachers, and have more than three years of experience in PE teaching and assessments. Eight teachers were selected, and their specific backgrounds are provided in Table 1.

Table 1. *Background information on participants*

Pseudonym	Age / Gender	Teaching career	Teaching type	Experience in PE assessments	Grade level	Region
Shin-young	30 / F	7 years	Classroom teacher	5 years	6th	City
Min-jeong	29 / F	6 years	Classroom teacher	5 years	5th	Rural
Tae-ho	29 / M	4 years	Specialist	4 years	4-6th	Rural
Ha-jin	41 / F	17 years	Classroom teacher	12 years	5th	Rural
Jong-soo	37 / M	11 years	Classroom teacher	10 years	6th	City
Jae-in	37 / M	11 years	Classroom teacher	8 years	6th	City
Woo-jin	33 / M	8 years	Specialist	8 years	5-6th	City
Soo-won	30 / M	5 years	Specialist	5 years	5-6th	City

Data Collection and Analysis

On-site data were collected, including with in-depth interviews, documents, photos, and video files. In-depth interviews using semi-structured questionnaires initially lasted 60 minutes. The following questions were contained in the semi-structured questionnaire, which covered participants' specific tasks, their knowledge of PE performance assessment and their source of information, and their critical thinking about PE performance assessment: "1. What kinds of tasks do you specifically conduct during your PE performance assessment?"; "2. What efforts have you been making to enhance your expertise regarding PE assessment?"; "3. What prerequisites would your ideal PE performance assessment need?"; and "4. What difficulties have you been through during your assessment of PE?" After the initial coding, additional 40-minute interviews were conducted with four research participants whose responses generated meaningful codes relevant to the four assessment literacy concepts. The additional interviews were based on detailed questions that examined the research participants' impartiality in the performance assessments, application of the performance assessment outcome, and teachers' efforts to develop students' assessment literacy.

As for on-site documents, data on the performance evaluation process for PE that research participants had conducted were elicited from the relevant documents and the resulting reports from previous PE performance assessments. Furthermore, the photos and video footage filmed during the participants' assessment of their students' PE performance were utilized since those data, though limited and indirectly confirming their on-site assessment, could be compared with what was stated in interviews.

The collected database was analyzed in the following stages, pursuant to an inductive categorical analysis as put forth by Patton (2002). By comparatively analyzing the collected on-site documents including research participants' plans for PE performance assessments, photos, and video footage, we obtained a comprehensive understanding of the PE performance assessment process. The theoretical framework consisting of the four conceptual factors of PE assessment literacy (Hay and Penney, 2013) was excellent in initiating a "start list" of pre-set codes (often referred to as "a priori codes"). However, the possibility of deriving other codes was not precluded in the coding process. Two coders worked independently for the initial coding, subsequently working on discordant codes again until they reached an agreement. The data analysis of in-depth interviews started with writing analytic memos after transcribing the interviews on the same day, from which initial codes were created by collecting repeated ideas during the segmenting task process. Thereafter, in-depth coding was conducted to create new inclusive codes by repetitively re-categorizing the initial codes. Finally, with a focus on the four conceptual factors of assessment literacy, outcomes were determined by inducing the themes that were interconnected among the created codes.

Using peer debriefing, we shared the initial data analysis with research participants to enhance the research veracity, performed checks to ensure data integrity, and also shared the entire research process with a university professor who had majored in Sports Pedagogy as well as three on-site teachers to guard against any error in the research process (Creswell, 2009). The entire process took place in accordance with the permission and regulations of the institutional review board (IRB) of the university with which the researcher of this study is affiliated.

Results and Discussion

Through the data analysis process, four themes were constructed that aligned with the four conceptual factors of PE assessment literacy (Hay and Penney, 2013): (i) lack of clarity about the concept of assessment, (ii) labor exchange and recycling, (iii) a lack of feedback, and (iv) teachers' critical thinking without pedagogy.

Assessment Comprehension: Lack of clarity about the concept of assessment

Assessment comprehension signifies teachers' knowledge and understanding about assessments in educational contexts. As a fundamental premise for assessment efficacy (Hay & Penney, 2009), assessment comprehension refers to how much an individual teacher is equipped with expertise and understanding of assessments and how they decide on the overall process and results of assessments; it is a starting point that determines the extent of influence the assessment results have on students.

Sufficient understanding of assessments should be preceded by an understanding of the concept of "assessment." However, the research participants exhibited confusion about assessment, evaluation, measurement, and their actual realization. Regarding this, Shin-young, one of the participants, stated:

> "Frankly, I am not quite sure what 'assessment' means. I am more familiar with the word 'evaluation,' and yet, at some point, the word 'assessment' started being used as well. As a teacher, I think PE performance assessments are conducted to evaluate, and classify into grades, students' physical functions and affective aspects related to classes. Is there much difference between 'evaluation' and 'assessment'?" (Shin-young)

What Shin-young stated regarding the conceptual definition of performance assessment still borders on a measurement-oriented assessment, which has long been conducted in PE classes. Other participants also could not transcend a peripheral procedure for assessments, such as measurement and grading. Such a lack of understanding of assessments led to a complete misunderstanding of performance assessment, which has led to a culture that emphasized measurement-oriented evaluation even after the introduction of performance assessments.

> "I was really dumbfounded when the performance assessment was first introduced. Similarly-nuanced different terms like performance assessment, authentic assessment, or alternative assessment kept pouring into the fields of education. There's not much difference compared to previous PE performance assessments, since these started without accurate instructions on how to conduct the assessment." (Tae-ho)

> "Those modifiers in front of the word 'assessment,' like course-oriented, comprehensive, or consistent, put quite a lot of pressure on teachers. They delineate the concepts but it would take significant efforts to have a clear understanding as to how to assess. Most realistically, teacher training would be a measure for upgrading their expertise, and yet PE training has not been properly conducted. Despite its name being PE assessment training, most of the training was practically not much more than just a one-time event, where ideas for class contents were shared." (Woo-jin)

The interviews with Tae-ho and Woo-jin revealed that the performance assessment system was initiated without an adequate introduction of the concept of performance assessment, and that insufficient training for currently working teachers was provided. Common issues were identified in that not enough opportunities were available for the PE teachers compared to those teaching knowledge-based subjects, and the training programs currently provided are decontextualized, one-time day training courses into which comprehensive instruction-oriented PE contents are poured (Armour & Yelling, 2004). Such problems in teacher training have dissuaded teachers from developing their assessment literacy. Therefore, teachers have conducted "instruction without assessments" and have rated their expertise on

assessment lower than that on instruction. In this regard, Soo-won stated:

> "Students' parents have less interest in PE assessments compared to their interest in knowledge-based subjects. Rather, I think that designing a PE class in a more fun and safe manner would better meet the needs of students and their parents. In this respect, I believe class expertise is more important than assessment expertise. Boring classes or classes where accidents occur result in complaints, but no complaints have been raised related to PE assessments." (Soo-won)

As seen in the above interview, Soo-won regarded instruction and assessments as mutually exclusive. This misconception is prevalent among teachers, and instruction and assessments are mutually interdependent. An assessment is a process that collects information not only about students' learning behaviors but also about teachers' teaching behaviors (McMillan, 2007; McMillan & Workman, 1999). Therefore, a good assessment process can enhance teachers' teaching abilities. Although teachers are only concerned with instruction for their own classes, conducting correct assessment based on their assessment literacy would increase their instruction capabilities. Furthermore, in the context of South Korea, where national-level physical education curriculum is specified in terms of its design, planning and evaluation the interaction between instruction (or pedagogies) and forms of assessments is strong. This is because the fit-for-curriculum assessments that actively incorporate what the national education program intends to achieve require a close-knit alignment between education course, class, and assessment (Penney et al., 2009). Consequently, it would be possible to enhance the efficiency of an education course if teachers' expertise, including assessment literacy, is properly developed.

Assessment Application: Labor exchange and recycling

Assessment application means that teachers' knowledge about effective assessment is formed in a classroom, through processes that help collect evidence for assessment interpretation (Hay & Penney, 2013). Generally, assessment application begins with formulating an assessment plan. However, research participants revealed a "labor-exchange and recycling" culture in the assessment planning phase, in which teachers share their work with others, as well as recycling that which is used in other classes. Ha-jin and Jong-soo said:

> "Except for two to three main subjects that specialist teachers teach, higher grade lead teachers generally need to come up with their assessment plan for eight to nine subjects. Too many. For this reason, lead teachers who teach the same grade divide their subjects and plan for their assessments. It's sort of like 'two hands are better than one.' Since they are assigned subjects based on their expertise, they are able to plan and implement quality assessments. Wouldn't it be improbable for us to be experts in all subjects?" (Ha-jin)

> "When planning for their performance assessments, many teachers either refer to, or use the entirety of the previous year's assessment plan, or plans shared on teachers' online communities. Good for saving time needed for planning, and reliable since those plans are made by teachers with expertise in PE." (Jong-soo)

As indicated above, it is not necessarily wrong for those with expertise on specific subjects to map out the performance assessment on those subjects and for teachers to share quality assessment plans. As seen in research participants' interviews, this could lead to a self-developing culture arising from a long-standing trial-and-error process led by the communities of elementary school teachers who need to teach various subjects by themselves. Planning a performance assessment should be preceded by a process that considers students' developmental levels, resources available, and consultations with teachers. All the previous performance assessments are the byproducts of such a process, from the previous year, and so some would consider them to be verified assessment plans, even though they do not take into account the current students. Furthermore, assessment plans shared through online teachers' communities often elicit replies from teachers who have actually implemented those plans, thus facilitating reduction of trial and error based on the simulations of the implementation process and results. These qualities will help utilize assessment plans in a labor-exchange and recycling manner.

However, such a labor-exchange and recycling style assessment culture has several issues. First, class assessments should be conducted in a class-specific manner. As for classroom assessments, tasks based on the contexts of students' actual life should be assigned (MacMillan, 2007; Baron, 1995), and for this to happen, aspects such as learners' starting behaviors, classroom ecological aspects, and available resources must be considered. A performance assessment planned in a labor-exchange fashion has issues that are more likely to induce identical assessment contents, methods, and procedures from all classes of each grade. A recycling style assessment also ignores various internal and external circumstances of classes, such as the revision of educational processes and changes in learners' demands.

Second, such a labor-exchange and recycling style assessment culture indicates someone else's assessment literacy, not that particular teacher's. With more teachers majoring in Sports Pedagogy, the quality of teacher guide books and other learning material has improved. In addition, the culture of online teachers' communities, where they share feedback on materials has assumed the form of peer scholarship. Nevertheless, it is doubtful whether sharing these assessment materials would be conducive to teachers' developing their own assessment literacy. It is important to identify whether a direct and holistic coverage of all the assessment procedures, ranging from planning to self-reflection relating to assessments, or indirect and partial involvement through sharing materials, would be more effective for developing teachers' assessment literacy.

Assessment Interpretation: A lack of feedback

Assessment interpretation is the process of interpreting collected information during the application of the assessment. In this procedure, it is necessary for teachers

to be equipped with assessment literacy, to determine their students' levels, and to chart directions for their future engagement and a groundwork for appropriate curricular and pedagogical adjustments. The results of assessment interpretation are transmitted to the students as feedback, so it is very important that the teacher's feedback includes these cues and that it is provided at the appropriate time.

Data analysis reveals that there has been a pervasive tradition among research participants to conduct an assessment during the last period of a unit and examine their assessment results at the end of the semester, which does not give them time to give feedback to the students. In this regard, research participants Shin-young and Tae-ho reported as follows:

> "If 12 T-ball classes are scheduled, a performance assessment will be conducted in the final (12th) class. This is because assessing during the last class would allow me to accurately assess which stage a student has reached." (Shin-young)

> "I usually leave the examination of assessment results to the semester's end. Based on my grading students on the day of the performance assessments, their final grades are decided and texts composed of four or five sentences are produced. On the last day of the semester, students and their parents are notified." (Tae-ho)

The analysis on other research participants' plans regarding performance assessments showed little difference in that they also measured assessment contents during the last unit, though with a bit of a time difference, and that they wrote and notified the assessment results at the semester's end. This signifies that our education fields have not been able to transcend Tyler's viewpoints-based assessment method, which regards the education process, teaching-learning, and evaluation cycle from a linear and sequential perspective. The most serious problem in Tyler's style of the output-oriented summative assessment lies in focusing on the core objective of assessments, that is, "feedback." Conducting the evaluation at the end of the semester means that the feedback on evaluation results would be generated after the completion of the course, due to which the students would not have the opportunity to amalgamate feedback into their performance. Feedback would be most effective when presented on time (Capel & Whitehead, 2015). Timely and appropriate feedback enhances students' performance ability and contributes to teachers' deciding and making arrangements about students' future assignments, their teaching contents, and methodology. When the formative assessment perspectives from which teachers record students' performance and give feedback to their students are maintained at each unit, timely and appropriate feedback can be formed (Earl, 2003). An assessment can be influential when a teacher becomes an effective mentor, kindred guide, and accurate reporter (Wilson, 1996).

Critical Engagement with Assessment: Teachers' critical thinking without pedagogy

The main points of teachers' literacy in relation to critical PE assessments are to recognize the importance and repercussions of the assessments, and to challenge the naturalness of assessment practice, performances, and outcomes. Therefore, they must challenge existing assessment methods in PE and consider the ramifications of the sociocultural assessments.

The neutrality of assessments has long been a dilemma. Technically, assessments cannot be neutral. Once the decision as to "what to assess" has been made, its neutrality is lost. Thus, teachers should predict which student would benefit and which student would be negatively influenced, and then formulate alternative approaches. Research participants had the following concerns with regard to the neutrality of their assessments:

> "It often happens that students with better physical health exercise get better grades in their performance assessments. Unlike other subjects, PE should be a subject in which students are free from assessment and able to have fun. The last thing I would like to see would be kids with lower kinetic functions having low PE grades. Therefore, I would give my students relatively better grades than in other subjects." (Ha-jin)

> "I don't think it's right for students' socio-economic status to have any influence on PE, or even on knowledge-based subjects. That's why I am trying to give them a better score on assessments if they are not able to go to private sports clubs or have sports experience with their parents, just to be fairer and more considerate with them." (Jae-in)

It cannot be said that research participants' linking the subject characteristics of elementary school PE to being "fun," and their considerations for children of a lower socio-economic status ruin the neutrality of the assessment. This can be considered an awareness of the purpose of PE subjects and critical thinking from a sociocultural perspective. However, such issues are not being considered in a pedagogical manner. Issues on prerequisite learning are also raised, as there are students who have a higher starting level in all subjects. Conducting a "give-away" assessment under the pretexts of distinct characteristics of PE and sociocultural influencing factors may spawn the undesired consequence of dragging the status of PE lower.

More focus should be placed on "process-oriented assessments." Though not a part of assessment results, the focus should be on how much students have grown compared to their starting behaviors in an assessment process (Hortigueala et al., 2016). Nonetheless, research participants are trying to be considerate, measuring results strictly from the perspective of assessment of learning. A prerequisite for the genuine process-oriented assessment is assessment for learning. This means being pedagogically considerate from an assessment-for-learning perspective to prepare a class strategy for students who have no prior experience; this would help them participate in assessments of their enhanced performance.

Furthermore, based on their critical thinking about existing PE and performance assessments, research participants endeavored to remove themselves from the framework. What they commonly selected for alternative assessments were self-assessments and peer-assessments, that is, student-led assessments. On the one hand, self-assessments would allow for critical

thinking and self-reflective awareness about their learning (Gordon, 1992; Hill & Ruptic, 1994), and peer-assessments could enhance the connectivity between the assessment process and learning (Green, 1994), both of which are somewhat encouraging. On the other hand, research participants' self- and peer assessments have critical issues. Comparison of self-assessments and peer-assessments records and the final performance assessment outcome document indicated that the research participants did not incorporate students' self-assessments and peer-assessments into the final document. Min-jeong and Soo-won stated the following reasons:

> "I sometimes conduct self- and peer assessments because they have their own share of effectiveness, but it's hard to put my trust in what my students have assessed themselves. In class, was the person my teammate, and out of class, was the person my friend? These factors have a more significant influence in PE than in any other subject." (Min-jeong)
>
> "The issue of neutrality doesn't only happen in peer assessments. There's this case where students who are favorable to other students' assessments apply strict and harsh standards to their own assessments." (Soo-won)

Several studies raised concerns regarding students' PE assessments, particularly on issues of neutrality (Brennan, 2001; Freeman, 1995; Porter & Cleland, 1995). However, given the fact that research participants decided on their students' final grades with a consideration of students' physical abilities and socioeconomic status when conducting assessments on affective fields, and that they also favored those students who had actively participated in class, who helped prepare the class, and who had behaved well (Black & Dockrell, 1980; Frey & Schmitt, 2010), the issue of outwardly appearing neutrality plagues both teachers and students alike. In addition, the discrepancy between the reliability of teacher assessment and of student assessment, to which research participants are referring, is also pertinent to consider. Probably, this would not involve the difference between teachers and students, but the difference in what each group has understood regarding assessment literacy.

An alternative to remedy teachers' distrust in student-led assessments would be to make their students assessment-literate (Hay & Penney, 2013). For that to occur, teachers must endeavor to improve their students' assessment literacy. It is necessary to provide students with more opportunities for peer assessments and to involve them in the entire assessment process, from communicating the purpose of the assessments to the utilization of results. Porter and Cleland (1995) mentioned that, initially, students become emotional during their peer-assessments, but subsequently, they partake in the essence of assessments. In addition, Freeman (1995) and Brennan (2001) pointed out the importance of assessor training to increase the credibility of student-led assessments. Students' assessment literacy can be enhanced only by more frequent participation in assessments.

Conclusions and Suggestions

This study examines the issues related to PE assessments in elementary schools from an assessment literacy perspective. Four main issues rose to the surface (*a lack of understanding of performance assessment, recycling of previous assessments, a lack of feedback to the students, and Teachers' critical thinking without pedagogy*); this information adds evidence to support the need to develop assessment literacy of both teachers and students. Based on this information, I would like to present three ways to address these issues:

First, the educational paths for current teachers to develop elementary school teachers' assessment literacy should be diversified. PE training programs for those currently teaching in the field for elementary schools are quantitatively and qualitatively insufficient (Lawrence, 2003; Harris et al., 2012). In-service teacher education, geared towards enhancing expertise in assessment, has been scarce (Stinggins & Conklin, 1992; Nitko & Brookhart, 2011). Developing assessment literacy should not be solely dependent on a teacher's individual efforts. It is necessary to provide in-service teacher education that can share assessment-related professional techniques and provide consistent learning opportunities in the contexts of both schools and educational communities (O'Sullivan & Deglau, 2006; Pritchard & Marshall, 2002).

Second, students should be provided with a specific guideline that leads them to become active evaluators. Teachers should not be satisfied with students in PE classes who are simply "busy, happy and good" (Placek, 1983). When a student becomes an active assessor, the assessment itself can impart learning (Black & William, 2006; Earl, 2003). Thus, a guideline that can enhance students' assessment literacy through theoretical research and on-site studies should be implemented in the relevant fields.

Finally, it is essential to assess the assessment process. The research participants had an opportunity for self-reflection as managerial level teachers or managers in the phases where they were planning and handling results. However, to increase the quality of performance assessments, the entire assessment process should be evaluated from the perspective of being tailored for the curriculum, class, and instruction. Assessment of the assessment process would strengthen teachers' literacy and equip them with a self-reflective opportunity to determine what has been missing in their literacy.

References

Annerstedt, C. & Larsson, S. (2010). 'I have my own picture of what the demands are...': Grading in Swedish-problems of validity, comparability and fairness. *European Physical Education Review*, *16*(2), 97-115. doi:10.1177/1356336x10381299

Armour, K. M. & Yelling, M. R. (2004). Continuing professional development for experienced physical education teacher: towards effective provision. *Sport, Education and Society*, *9*(1), 95-114. doi:10.1080/1357332042000175836

Baron, M. (1995). *Authentic assessment: The key to unlocking student success*. Lancaster, PA: Techonomic.

Black, H. D. & Dockrell, W. B. (1980). Assessment in the Affective Domain. *British Journal of Educational Research*, *6*(2), 197-208. doi:10.1080/0141192800060208

Black, P. & William, D. (2006). Developing a theory of formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 81-100). London: Sage.

Brennan, R. L. (2001). Generalizability theory. New York: Springer.

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks, CA: Corwin.

Fullan, M. (2002). The role of leadership in the promotion of knowledge management in schools. *Teacher and Teaching: Theory and Practice*, *8*(3), 409-419. doi:10.1080/1354060021000005 30

Freeman, M. (1995). Peer assessment by group of groups work. *Assessment & Evaluation in Higher Education*, *20*, 289-300. doi:10.1080/0260293950200305

Frey, B. B. & Schmitt, V. L. (2010). Teachers' Classroom Assessment Practices. *Middle Grades Research Journal*, *5*(3), 107-117.

Gordon, M. J. (1992). Self-assessment programs and their implications for health professions training. *Academic Medicine*, *67*, 672–679. doi:10.1097/00001888-199210000-00012

Green, M. (1994). Children as evaluation understanding from the inside. In Cambourne, B. & Turbil, J. (Eds.). *Responsive Evaluation. Portsmouth* (pp. 83-103). NH : Heinemann.

Hardman, K. & Marshall, J. J. (2000). *World-wide survey of the state and status of school physical education*. Manchester: University of Manchester press.

Harris, J., Cale, L., & Musson, H. (2012). The predicament of primary physical education: a consequence of 'insufficient' ITT and 'ineffective' CPD? *Physical Education and Sport Pedagogy*, *17*(4), 367-381. doi:10.1080/17408989.2011.582489

Hay, P. & Penney, D. (2009). Proposing conditions for assessment efficacy in physical education. *European Physical Education Review*, *15*(3), 389-405. doi:10.1177/1356336x09364294

Hay, P. & Penney, D. (2013). *Assessment in physical education: A sociocultural perspective*. London: Routledge.

Hill, B.C., & Ruptic, C. (1994). *Practical aspects of authentic assessment: Putting the pieces together*. Norwood, MA: Christopher-Gordon.

Hortigueala, D., Fernandez-Rio, J., & Perez-Pueyo, A. (2016). Long-term effects of the pedagogical approach on the perceptions of physical education by students and teachers. *Journal of Physical Education and Sport*, *16*(4), 1326-1333. doi:10.7752/jpes.2016.04210

Lawrence, J. (2003). The impact and use of the TOPs programme in one local education authority. *British Journal of Teaching Physical Education*, *34*, 44-48.

Matanin, M. & Tennehill, D. (1994). Assessment and grading in physical education. *Journal of Teaching in Physical Education*, *13*(4), 395-405. doi:10.1123/jtpe.13.4.395

McMillan, J. H. (2007). *Classroom assessment: Principles and practice for effective standards-based instruction* (4th ed.). Boston: Allyn & Bacon.

McMillan, J. H. & Workman, D. (1999). *Teachers' classroom assessment and grading practices: Phase 1*. Richmond, VA: Metropolitan Educational Consortium.

Nitko, A. J. & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). MA: Pearson.

O'Sullivan, M. & Deglau, D. (2006). Principles of professional development. *Journal of Teaching in Physical Education*, *25*, 441-449. doi:10.1123/jtpe.25.4.441

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Penney, D., Brooker, R. Hay p., & Gillespie L. (2009). Curriculum, pedagogy and assessment: three message systems of schooling and dimensions of quality physical education. *Sport, Education and Society*, *14*(4), 421-442. doi:10.1080/13573320903217125

Placek, J. H. (1983). Conceptions of success in teaching: busy, happy, and good? In Templin, J & Olsen, J. (Eds.). *Teaching in Physical Education* (pp.45-56). Champaign, IL: Human Kinetics.

Porter, C. & Cleland, J. (1995). *The portfolio as a learning strategy*. Portsmouth, NH: Boynton/Cook.

Pritchard, R. J. & Marshall, J. C. (2002). Professional development in 'healthy' vs. 'unhealthy' districts: top ten characteristics based on research. *School Leadership and Management*, *22*, 113-141.

Stiggins, R. J. & Conklin, N. F. (1992). *In teachers ` bands: Investigating the practices of classroom assessment*. Albany, NY: SUNY.

Thompson, M. D., & Penney, D. (2015). Assessment literacy in primary physical education. *European Physical Education Review*, *21*(4), 485-503. doi:10.1177/1356336x15584087

Wilson, R. (1996). Assessing students in classrooms and schools. Toronto: Allyn & Bacon.

Yoo, J. (2005). Understanding the developing process of physical education curriculum at the secondary physical educator level. *Korean Journal of Sport Pedagogy*, *12*(1), 31-56.

This page is intentionally left blank

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effect of Compensation Studies on Disadvantaged Children's Bully Behaviours

Gönül Onur Sezer [a]

Received: 9 October 2017
Revised: 24 November 2017
Accepted: 28 November 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236117

Abstract

Children who are economically, educationally, linguistically or socially disadvantaged called "disadvantaged children". Those children are at risk and they must be supported because of their negative life conditions. Compensation studies must be implemented to those children. The "Bir Umut Ol Benim İçin" (Be My Hope) project is such a compensation study which is held by Uludag University Faculty of Education and Bursa Police Department Children Branch Office. The aim of this study is to determine the impact of this project on disadvantaged children's bullying behaviours. In line with this objective Colorado School Climate Survey, which was developed by Garrity et al., (2000) was used. The questionnaire was designed to measure several aspects of bullying. The analyses of the study showed that the project "Bir Umut Ol Benim İçin" has not caused a positive effect on the bully behaviours. While it was hoped that this project would cause a decrease in bully behaviours of disadvantaged children because it is a compensation study, it is seen that such as studies/projects must be supported with bully proofing and with conflict resolution programs.

Keywords: Disadvantaged children, bully behaviour, compensation study, risk factor.

Introduction

Some children are economically, educationally, linguistically or socially disadvantaged. They lack of the basic necessities of life, they have been denied the basic and universal rights of children, the opportunity to grow normally at their own natural rate. The term of "disadvantaged children" is used to specify those children who are subjected to detrimental environmental stresses of any kind and handicapped or disabled because of certain conditions of exogenous risks and lastly, who are at risk of future psycho-educational problems (Dash, 2007; Moore, 2006). Schatz, Smith, Borkowski, Whitman and Keogh (2008) stated that, ill-treated children are more capable of showing withdrawn behaviours, social problems and bully behaviours, get lower grades and fail the class. Supportive studies such as compensative studies must be implemented for those children. Compensation studies which require a sophisticated and comprehensive study consist of activities aiming to overcome the problems created by existing risk factor (Kırcaali İftar, 2007).

Compensation studies help disadvantaged children to overcome this disharmony social organization (Smith et al., 2004). These compensation studies aim to help the children readjust and require a suitable education and training program that is prepared with pedagogic and scientific approach by determining their intelligence, characteristics, talents, and closely observing his development stage. Thereby, it is intended to help the child develop his lacking behaviours and become an individual that can obey social rules (Sarpdağ, 2005). The studies about bullying started in later 1970s, especially in Norway, Sweden and Finland. Particularly, "Bullying in Schools", the book published in 1978 by the Norwegian researcher Dan Olweus, is taken as a milestone in bullying. Recently these studies have increased in England, Australia and the USA. While there is not a universal definition about bullying, researchers generally defines bullying as repetitive bully behaviours, asymmetric power relations or systematic abuse of power (Olweus, 1999). Some researchers identified the bullying as intentional bully behaviours against others; some others state that, in order to consider these kinds of behaviours as bullying, they have to be repeated regularly (Griffin & Gross, 2004).

The commonly used definition belongs to Olweus (1987): "If a person is confronted with intentional, repetitive and continuing negative behaviours at least for a while by another person or persons, it can be said that this person is sustained from bullying". Negative behaviours are explained as behaviours that hurt and bother someone intentionally or the behaviours that try to do this. This definition enables us to differentiate the random bully behaviour from bullying (Griffin & Gross, 2004). Dominant person(s) harm(s) the less dominant people intentionally and repetitively. This bully behaviour may be direct or indirect, both physically (for example; kick, hit) or orally (for example; nicknaming). Direct bullying occurs by definite bullying such as hitting, pushing, japing, menacing and damaging the properties that the other person had. On the other hand, indirect bullying which is carried out via behaviours that do not necessitate the person who bullies and the bully victim to meet directly such as

[a] Correspondence details: Gönül Onur Sezer, Uludag University, Faculty of Education, Department of Elementary Education, Room Number: 206, Gorukle Campuss, 16059, Bursa, Turkey. E-Mail: gonulonur@uludag.edu.tr

excluding from the social group, spreading rumours and humiliating. By this way the status and sense of the victim is changed by the others. This kind of bullying is also detrimental at least as much as direct bullying; however it has entered the field of psychosocial research recently (Connor, 1998).

Bullying causes serious problems for victims and bullies. Broken down of the school climate and the training of children in a safe environment because of bullying in short term, may cause toddling off from the school, degeneration of psycho-social conformity, criminal behaviours and apprehension, loneliness and depression for the victims in the long term (Scarpaci, 2006). The most important long term effect of bullying is the misuse of intensity and power. A high correlation is appointed between the bully behaviours in primary school and crime rate that arise in the upcoming years (Olweus, 1993). At the same time, the tendency of victims about their self-confidence to be lower at adulthood (Byrne, 1994). Ersanlı (2007) states that punishment has been considered to be the only method for authority and deterrence up until today but studies such as compensation studies, towards reserializing and deterrence have gained importance in recent years. The aim of the project "Be My Hope", which was applied in Bursa during spring semester of the year of 2011-2014 and carried out by volunteer teacher candidates and children under the risk in cooperation with the Children Branch Office of Bursa Police Department and the Faculty of Education of Uludağ University, is to provide that primary education students in the risk group who cannot benefit from equal opportunities like students of the same age because of unfavourable conditions of the environment where they live, conduct academic and social activities with teacher candidates who receive education in the Faculty of education. In addition teacher candidates become role models and that these children are reintegrated into the society as individuals who have respect for themselves and environment, who are happy, active and productive. Another purpose of the project is to perform academic and social activities to the children who are under risk and cannot have the equal opportunities with their peers due to the unfavourable environment they have lived with teacher candidates. In this way these children would be reintegrated into the society by being self and environment respecting, happy, effective and productive individuals. At the same time during the project the teacher candidates would be the role model for those children. Within the context of the project, the academic studies, social and sports activities are organized for the academic and social development of the children under risk.

The Aim of the Study

"Bir Umut Ol Benim İçin" project is a study that encourages the academic and social development of children under risk, to reintegrate into the society as healthy, happy and productive individuals. From this point forth, the aim of this study is to examine the effect of the study of compensating the risk factors on the bullying behaviours of children, who were qualified as disadvantaged, within the scope of the project. Therefore this study is important for being a research to the body of literature.

Methodology

For the purpose of the project, children chosen by Bursa Provincial Security Directorate's Child Branch benefit from this process by participating individually or as a group in academic and social activities with volunteer teacher candidates during planned days and hours. Within the scope of this Project study times, social activities and sport activities are held for their academic developments by teacher candidates and disadvantaged children in recreational and sports facilities provided by Metropolitan Municipality of Bursa. Teacher candidates organize study times for academic development of disadvantaged children groups in environments provided by City Council. In these studies, Bursa Provincial Security Directorate's Child Branch locates disadvantaged children, obtains the consent of families, transports children to related study areas and provides security during the process of studies. The explanations about the model of the study, study group, data collection tools and statistical techniques that were used in the analysis of the collected data are included in this section.

Research Model

This research is carried out in conformity within the scope of a project and the results of the pre-test and post-test were examined, which aims to test the significant difference between the related two surveys. It is examined in the study that whether there were significant differences between the pre-test and post-test scores of Colorado School Climate Survey (Garrity, Jens, Porter, Sager & Short-Camilli, 2000) including the statements for the determination of bully behaviours of the children who had attended the study.

Participants

The participants of the study consisted of 33 children, who were evaluated as disadvantaged by Bursa Police Department Children Branch Office and who were included in "Bir Umut Ol Benim İçin" project. The information about the children, who attended the study, is shown in Table 1.

Table 1. *Descriptive Statistics about the Children Who Attended the Project*

		N	*%*
Gender	Girl	5	15.2
	Boy	28	84.8
Total		33	100
Class	5. Grade	1	3
	6. Grade	25	75.8
	7. Grade	3	9.1
	8. Grade	4	12.1
Total		33	100
Age	11	1	3
	12	13	39.4
	13	14	42.4
	14	5	15.2
Total		33	100

When Table 1 is examined, it is seen that 33 children 5 of which (15.2%) are girls and 28 of which (84.8%) are boys. 42.2 % of the children are 13; 39.4% are 12; 15.2% are 14 and 3% are 11 years old. 75.8% of the children are in 6. grade, 12.1% to 8. grade, 9.1% to 7. grade and 3% to 5. grade.

Data Gathering Tools

In this study the Colorado School Climate Survey, which was developed by Garrity et al, as a part of a bully proofing program for schools was used. (2000). The Turkish form of this questionnaire with some small modifications has been done by three educational psychologists Kartal and Bilgin (2007; 2008; 2009). The questionnaire form consists of the statements about who had done the bullying, the places where they suffered from bullying, to whom they had explained the bullying. The questionnaire was 4-point Likert type. In this study, in the beginning of the project the Cronbach Alpha coefficient was determined as .62 and in the end of the project it was found as .75.

Data Analysis

The pre-test and post test scores obtained from the questionnaire was given in findings section by being analysed with chi-square (X^2) test and the results were achieved after they have been interpreted in later sections. In the analysis of the data, the difference between student pre-test and post-test reports in terms of the proportions of bullying types and bully Kay-Square test was carried out to determine if it was significant. The significance value was set at .05.

Results

When Table 2 is examined, a significant difference is identified between the pre-test and post-test percentages of the answers to the item *"Other children hit or kicked me"* [$X^2_{(9)}$= 20.621, p<.05]. An increase is observed in the post test percentages of the children who said "Never" and a decrease is observed in the percentages of the children who said "once a week", "2-4 times a week" and "more than 5 times a week".

Table 2. *The Chi-square Test Results about the Comparison of the Pre-Test and Post-Test for the Bully Behaviours the Children Suffer From and Their Frequency*

During the last month		Pre-test (%)	Post-test (%)	*p*
Other children hit or kicked me	Never	51.5	66.7	.014*
	Once a week	33.3	24.2	
	2-4 times a week	6.1	3.0	
	More than 5 times a week	9.1	6.1	
Other children spoke ill of me, ridiculed me or nicknamed me	Never	62.5	53.1	.010*
	Once a week	21.9	28.1	
	2-4 times a week	6.3	6.3	
	More than 5 times a week	9.4	12.5	
Other children did not let me join them	Never	78.8	81.8	.000*
	Once a week	15.2	9.1	
	2-4 times a week	3.0	3.0	
	More than 5 times a week	3.0	6.1	
Other children took the things that belonged to me.	Never	56.3	78.1	.001*
	Once a week	28.1	12.5	
	2-4 times a week	12.5	3.1	
	More than 5 times a week	3.1	6.3	
Other children threatened me to hurt me and to take my objects.	Never	83.9	83.9	.858
	Once a week	9.7	3.2	
	2-4 times a week	0	9.7	
	More than 5 times a week	6.5	3.2	

*p<.05

A significant difference is also identified between the pre-test and post-test percentages of the answers to the item *"Other children spoke ill of me, ridiculed me or nicknamed me"* [$X^2_{(9)}$= 21.630, p<.05]. When the answers of the children are examined, while a decrease is being observed in the post test percentages of the children who said "Never", there has been an increase in the post test percentages of the children who said "once a week" and "more than 5 times a week". On the other hand, there has been no change in the test percentages of the children who said "2-4 times a week".

A significant difference is also identified between the pre-test and post-test percentages of the answers to the item *"Other children did not let me join them"* [$X^2_{(9)}$= 33.940, p<.05]. An increase is determined in the post test percentages of the children who said "Never" and "more than 5 times a week"; a decrease is determined in the post test percentages of the children who said "once a week". On the other hand, no change has been observed in the test percentages of the children who said "2-4 times a week".

A significant difference is determined between the pre-test and post-test percentages of the answers to the item *"Other children took the things that belonged to me"*. [$X^2_{(9)}$= 29.431, p<.05]. When the answers are examined, an increase in the post test percentages of the children who said "Never" and "more than 5 times a week" and a decrease in the post test percentages of the children who said "once a week" and "2-4 times a week".

There is not a significant difference between the pre-test and post-test percentages of the answers to the item *"Other children threatened to hurt me and to take my objects"* [$X^2_{(6)}$= 2.588, p>.05]. While a decrease is seen in the post test percentages of the children, who said "once a week" and "more than 5 times a week", an increase is seen in the post test percentages of the children who said "2-4 times a week". No change is observed in the test percentages of the children who said "Never".

Table 3. *The Chi-square Test Results about the Comparison of the Pre-Test and Post-Test for the Places Where the Children Suffer From Bullying*

Place Where I Suffer From Bullying	Pre-test (%)	Post-test (%)	p
Classroom	47.6	33.3	
Garden	0	23.8	
Hallway	9.5	19.0	
Toilet	4.8	9.5	.484
Round trip road to school	28.6	9.5	
Canteen	0	4.8	
Nowhere	9.5	0	

The pre-test shows that the children mostly (47.6%) have suffered from bullying in classrooms, then respectively *"the round trip road to school"* (28.6%), *"hallway"* (9.5%) and *"toilet"* (4.8%). Children left *"the canteen"* and *"the garden"* options empty. 9.5% of the children gave the answer *"nowhere"* to this question. When the post test results are examined, though there has been a decrease in the percentages of the children, who has given the answer *"classroom"* (33.3%), even so it is determined that they have suffered from bullying in the class at most. *"Garden"* (23.8%) and *"hallway"* (19.0%) follow this answer. It is seen that the post-test percentages of "toilet" and *"round trip road to school"* is the same as 9.5%. *"Canteen"* is at last with 4.8%. According to the answers given by the children, it is seen that there has not been a significant difference between the pre-test and post-tests [$X^2_{(20)}$= 19.580, p>.05].

Table 4. *The Chi-square Test Results about the Comparison of the Pre-test and Post-tests for by Whom the Children Were Bullied*

The Bully	Pre-test (%)	Post-test (%)	p
A girl in the same class with me	0	10.5	
A boy in the same class with me	36.8	31.6	
A girl in an upper grade	0	5.3	
A boy in an upper grade	21.1	42.1	.274
Girls in the same class with me	5.3	0	
Boys in the same class with me	31.6	10.5	
None	5.3	0	

In the pre-test for the determination of by whom the children were bullied, it is seen that the item "*A boy in the same class with me*" has the highest percentage (36.8%) and the item "*Boys in the same class with me*" (31.6%) follows this answer. The percentage of the third item "*A boy in an upper grade*" is 21.1%. While it is determined that 5.3% of the children have been bullied by the girls from the same class with themselves, it is identified that 5.3% of the children have answered the question as "none". When the post test is examined, it is seen that the item *"A boy in an upper grade"* has the highest percentage (42.1%) and the followed item is *"A boy in the same class with me"* (31.6%). While it is determined that the percentages of the items *"A girl in the same class with me"* and "*Boys in the same class with me"* have remained the same (10.5%), it is identified that the percentage of the ones that have given the item *"A girl in an upper grade"* has been 5.3%. It is seen that there has not been a significant difference between the pre-test and post-tests percentage to the items of the children [$X^2_{(16)}$= 18.906, p>.05].

Table 5. *The Chi-square Test Results About The Comparison of The Pre-Test and Post-Tests for The People to Whom The Children Told About The Bullying*

To whom did you tell the bullying?	Pre-test (%)	Post-test (%)	p
Nobody	45.5	59.1	
To a friend	9.1	22.7	.691
To an adult at school	18.2	4.5	
To mother-father	27.3	13.6	

When the people to whom the children had told about the bullying are examined, while it is seen that the ones, who have given the answer *"Nobody"*, were 45.5% and the ones who had told to their mothers and fathers have followed this answer with 27.3% in the percentages of the pre-test. 18.2% percent of the children told about the

bullying that they had suffered to an adult at school and 9.1% told to a friend. The ones who have given the answer *"Nobody"* have been the first (59.1%) and this had been followed by the ones who had given the answer *"to a friend"* with 22.7% at post-test. While the percentage of the ones, who had told about it to their mothers and fathers were 13.6%, it is observed that there is a decrease in the percentage of the ones who had told an adult at school (4.5%). It is identified that there is not a significant difference between the pre-test and post-test according to the answers of the children [$X^2_{(9)}$= 6.480, p>.05].

Table 6. *The Chi-square Test Results about the Comparison of the Pre-Test and Post Tests for the Bully Behaviours of the Children and Their Frequency*

In this period		Pre-test (%)	Post-test (%)	*p*
I hit or kicked the other children	Never	51.5	42.4	.944
	Once a week	33.3	24.2	
	2-4 times a week	12.1	15.2	
	More than 5 times a week	3.0	18.2	
I spoke ill of the other children, ridiculed them or nicknamed them	Never	64.5	58.1	.803
	Once a week	19.4	22.6	
	2-4 times a week	12.9	12.9	
	More than 5 times a week	3.2	6.5	
We did not let the other children join us	Never	81.8	72.7	.382
	Once a week	6.1	21.2	
	2-4 times a week	9.1	6.1	
	More than 5 times a week	3.0	0	
I took the objects that belong to other children.	Never	90.6	90.6	.133
	Once a week	6.3	3.1	
	2-4 times a week	3.1	6.3	
	More than 5 times a week	0	0	
I threatened the other children to hurt them or to take their objects.	Never	93.5	80.6	.916
	Once a week	6.5	6.5	
	2-4 times a week	0	9.7	
	More than 5 times a week	0	3.2	

There is no significant difference between the pre-test and post-test percentages of the answers given to the item *"I hit or kicked the other children"* [$X^2_{(9)}$= 3.451, p>.05]. A decrease is observed in the percentages of the children who said "Never" and "once a week"; an increase in observed in the percentages of the children who said "2-4 times a week" and "more than 5 times a week".

A significant difference is not determined between the pre-test and post-test percentages of the answers given to the item *"I spoke ill of the other children, ridiculed them or nicknamed them"* as well [$X^2_{(9)}$= 5.345, p>.05]. When the answers given by the children are examined, while a decrease is being observed in the post test percentages of the children, who said "Never", there was an increase in the post test percentages of the children who said "once a week" and "more than 5 times a week". On the other hand, there was no change in the post test percentages of the children who said "2-4 times a week".

A significant difference is not identified between the pre-test and post-test percentages of the answers given to the item "*We did not let the other children join us*", as well [$X^2_{(6)}$= 6.384, p>.05]. A decrease in the post test percentages of the children, who said "Never", "2-4 times week" and "more than 5 times a week" and an increase in the post test percentages of the children who said "once a week" is determined.

No significant difference is identified in the pre-test and post-test percentages of the answers given to the item *"I took the objects that belong to other children"* [$X^2_{(4)}$= 7.049, p>.05]. No change is observed in the percentages between the answers "Never" and "more than 5 times a week" in the pre-test and post-test. On the contrary, it is seen that there is a decrease in the post test percentages of the children who said "once a week" and an increase in the post test percentages of the children who said "2-4 times a week".

There is a significant difference between the pre-test and post-test percentages of the answers given to the item *"I threatened the other children to hurt them or to take their objects"* [$X^2_{(3)}$= .513, p<.05]. A decrease in the post test percentages of the children who said "Never" and an increase in the post test percentages of the children who said "2-4 times a week" and "more than 5 times a week" are mentioned. There is no change in the post test percentage of the children who said "once a week".

Table 7. *The Chi-Square Test Results About The Comparison of The Pre-Test and Post-Tests For The People Who Were Bullied by Children*

The one whom I bully	Pre-test (%)	Post-test (%)	*p*
In the same class with me	61.9	61.9	.488
In a lower grade	9.5	4.8	

In an upper grade	19.0	33.3	
None	9.5	0	

When Table 7 is examined, it is determined that the children had mostly bullying the children, who were in the same class with them and these children have the same percentages in both test applications (61.9%). An increase (33.3%) in the post test percentages of the children who said *"In an upper grade"* and a decrease (4.8%) in the post test percentages of the children who said *"In a lower grade"* is observed. No significant difference is identified between the pre-test and post-test applications about this article [$X^2_{(6)}$= 5.450, p>.05].

Table 8. *The Chi-square Test Results about the Comparison of the Pre-Test and Post-Tests for the Places Where the Children Bully*

Place where I bully	Pre-test (%)	Post-test (%)	*p*
Classroom	50.0	22.7	
Garden	13.6	22.7	
Hallway	4.5	31.8	
Toilet	9.1	13.6	.605
Round trip road to school	13.6	9.1	
Nowhere	4.5	0	
All of them	4.5	0	

The pre-test results of the children indicate that the place where they bully mostly had been the class (50.0%), then respectively the garden (13.6%), round trip road to school (13.6%) and toilet (9.1%). 4,5% of the children said *"none, hallway and all of them"* to this question. When the answers of the post test is examined, a decrease (22.7%) is seen in the percentages of the children, who said that the place where they bully was *"the class"* and a serious increase is seen (31.8%) in the percentages of the children who said that the place where they bully had been *"the hallway"*. "*The classroom"* and *"the garden"* is following the *"the hallway"* with 22.7%. It is determined that the ones who said *"the toilet"* is 13.6% and the ones who said *"round trip road to school"* is 9.1%. "None" and "All of them" is at last with 4.8%. They left "none" and "all of them" options empty. It is seen that there is no significant difference between the pre-test and post-tests according to the answers given by the children [$X^2_{(24)}$= 21.562, p>.05].

Discussion, Conclusions and Recommendations

When it is looked at the frequencies of the bully behaviours that the children suffer from, while a significant difference is being determined in favour of the post-test in the items "*Other children hit or kicked me", "Other children did not let me join them" and "Other children took the things that belong to me"*, a significant difference is being determined in favour of the pre-test in the item *"Other children spoke ill of me, ridiculed me or nicknamed me".* On the contrary, no significant difference was identified in the article "*Other children threatened to hurt me or to take my objects".* It can be said that there has been a decrease in the bully behaviours the children suffer from in general.

The children stated in both tests that they have mostly been bullied in the classrooms. A similar result about the bullying of the children mostly in the class was also found by Özkan and Gökçearslan (2010). Again Kartal and Bilgin (2008) stated in their studies that the children mostly have been bullied in the classroom. However, Çınkır and Karaman Kepenekçi (2003) determined that the bullying mostly had occurred in the school garden and secondly in the classroom. In the analyses of this study, a significant difference in the people, who bully the children, is out of question. When the percentages are examined, the children stated in the pre-test that they have mostly been bullied by a boy in the same class with them and in the post test that they have mostly been bullied by a boy in an upper grade. Bullying of the children by the boys in both tests might be about the majority of the participants that constituted of mostly boys. In the study carried out by Özkan and Gökçearslan (2010), it was concluded that mostly boys were bullying. A similar conclusion was determined in the studies carried out by Çınkır and Karaman Kepenekçi (2003), Kartal and Bilgin (2008) and Kartal (2009).

It is determined that the people to whom the children have told about the bullying had not shown a significant difference and that they had preferred no one to tell this in both tests. This case may be interpreted as the children did not want to tell the bullying they suffer from to others in order not to be seen weak. In the study of Kartal and Bilgin (2008), the children stated that they had told about the bullying they suffer from to no one too.

There is no significant difference in the pre-test and post-tests of the bully behaviours of the children and their frequencies. In a qualitative study performed for taking the opinions of the same sample group about "Bir Umut Ol Benim İçin" project, a child stated that *"My behaviors have changed. I behave well to others. I used to involve in a fight before, now I wait for a while."* This opinion of the child shows that the project caused positive changes in bully behaviours of some children. Uysal (2006) identified that "Education Program against Violence" conducted with seventh grades had not caused a significant difference in the bully behaviour points of the children too. However no significant difference is determined in the people who were bullying by children and in the places where the children bully. The children stated that they have mostly bullying the people in the same class with them.

When all these results are evaluated, it is concluded that "Bir Umut Ol Benim İçin" project has not had a positive

effect on the bully behaviours of the children. Although it was hoped that the project contributed to disadvantaged children in this regard, the negative results obtained are evaluated as an expected situation since the process had not been a proof program from bullying. Eslea and Smith (1998) after the application of a school wide biennial program against bullying, stated that the number of children, who had said that they have been bullied, had decreased and in the next year while this situation remained the same in a school, it had begun to increase in another school. In the end of another application carried out by Stevens and his friends (2000), it was specified that though the frequency of suffering from bullying have remained the same, there had been significant decreases in the levels of bullying. Pepler and his friends (1994), in the end of their study, found that while the frequency of suffering from bullying was decreasing, on contrary the frequency of bullying others have increased. Epstein, Plog and Porter (2000), in the end of a four-year program, did not encounter a reduction in exclusion rates despite the reduction in physical and oral bullying (Beran & Tutty, 2002). However Uysal (2006) stated that the "Education Program against Violence" which he has conducted with seventh grades had increased the conflict resolution abilities of the children and had significantly decreased their tendency to violence. Hence it is thought that bully proofing programs must be supported with conflict resolution programs. Thus the child will be learning to find a solution to bullying that he suffered from with his conflict resolution abilities. In addition it was observed that "Awareness towards Violence" program conducted by Düzgün, Alibeyoğlu and Orhan (2006) had increased the awareness of the children for violence and violent behaviours and that the children have developed behaviours. These two studies show that the programs devoted to protect from bullying or violence have caused positive results.

From this point of view, it may be suggested that the studies that would be carried out with disadvantaged children to be performed by including the program of protection from bullying and the program of conflict resolution abilities. The study is a quantitative study and it is thought that the deep analysis of the results obtained would make significant contributions to the field. Accordingly, these kinds of studies may be supported with qualitative studies. Also teacher observations may be included. Because, the majority of the participants consisted of boys caused to obtain restricted data about the roles of the girls in bullying process. Performing these kinds of analyses with sample groups, in which the girls are at quorum, is another suggestion of the study. The long term success of the bullying protection programs is possible with the cooperation of teachers, school staff, family members and children.

Recommendations

When it is looked at the frequencies of the bully behaviors that the children suffer from, while a significant difference is being determined in favor of the post-test in the items "*Other children hit or kicked me", "Other children did not let me to join themselves" and "Other children took the things that belong to me"*, a significant difference is being determined in favor of the pre- test in the item *"Other children spoke ill of me, ridiculed me or nicknamed me".* On the contrary, no significant difference was identified in the article "*Other children threatened me to hurt me or to take my objects".* It can be said that there have been a decrease in the bully behaviors the children suffer from in general.

The children stated in both tests that they have mostly been bullying in the class. A similar result about the bullying of the children mostly in the class was also found by Özkan and Gökçearslan (2010). Again Kartal and Bilgin (2008) stated in their studies that the children mostly have been bullying in the class. However Çınkır and Karaman Kepenekçi (2003) determined that the bullying mostly had occurred in the school garden and secondly in the class. In the analyses of this study, a significant difference in the people who bully the children is out of question. When the percentages are examined, the children stated in the pre-test that they have mostly been bullying by a boy in the same class with them and in the post test that they have mostly been bullying by a boy in an upper class. Bullying of the children by the boys in both tests might be about the majority of the participants that constituted of mostly boys. In the study carried out by Özkan and Gökçearslan (2010), it was concluded that mostly boys had bullying. A similar conclusion was determined in the studies carried out by Çınkır and Karaman Kepenekçi (2003), Kartal and Bilgin (2008) and Kartal (2009).

It is determined that the people to whom the children have told the bullying had not shown a significant difference and that they had preferred no to tell this to anyone in both tests. This case may be interpreted as the children did not want to tell the bullying they suffer from to others in order to be seen weak. In the study of Kartal and Bilgin (2008), the children stated that they had told the bullying they suffer from to no one too.

There is no significant difference in the pre-test and post-tests of the bully behaviors of the children and their frequencies. In a qualitative study performed for taking the opinions of the same sample group about "Bir Umut Ol Benim İçin" project, a child stated that *"My behaviors have changed. I behave well against others…… I used to involve in a fight before, now I wait for a while."* This opinion of the child shows that the project caused positive changes in bully behaviors of some children. Uysal (2006) identified that "Education Program against Violence" conducted with seventh classes had not caused a significant difference in the bully behavior points of the children too. However no significant difference is determined in the people who were bullying by children and in the places where the children bully. The children stated that they have mostly bullying the people in the same class with them.

When all these results are evaluated, it is concluded that "Bir Umut Ol Benim İçin" project has not had a positive effect on the bully behaviors of the children. Although it was hoped that the project to have contributed to disadvantaged children in this regard, the negative results obtained are evaluated as an expected situation since the process had not been a proof program from bullying. Eslea and Smith (1998) after the application of a

schoolwide biennial program against bullying, stated that the number of children who had said that they have been bullying had decreased and in the next year while this situation have been remaining the same in a school, it had begun to increase in another school. In the end of another application carried out by Stevens and his friends (2000), it was specified that though the frequency of suffering from bullying have remained the same, there had been significant decreases in the levels of bullying. Pepler and his friends (1994), in the end of their study, found that while the frequency of suffering from bullying was decreasing, on contrary the frequency of bullying others have increased. Epstein, Plog and Porter (2000), in the end of a four-year program, did not encounter a reduction in exclusion rates despite the reduction in physical and oral bullying (Beran & Tutty, 2002). However Uysal (2006) stated that the "Education Program against Violence" which he has conducted with seventh classes had increased the conflict resolution abilities of the children and had significantly decreased their tendency to violence. Hence it is thought that bully proofing programs must be supported with conflict resolution programs. Thus the child will be learn to find a solution to bullying that he suffered from with his conflict resolution abilities. Again it was observed that "Awareness against Violence" program conducted by Düzgün, Alibeyoğlu and Orhan (2006) had increased the awareness of the children for violence and violent behaviors and that the children have developed behaviors that away after the program. These two studies show that the programs devoted to proof from bullying or violence have caused positive results.

From this point forth, it may be suggested that the studies that would be carried out with disadvantaged children to be performed by including the program of proof from bullying and the program of conflict resolution abilities. The said study is a quantitative study and it is thought that the deep analysis of the results obtained would have significant contributions to the field. Accordingly, these kinds of studies may be supported with qualitative studies. And teacher observations may be included. Since the majority of the participants are constituted of boys caused to obtained restricted data about the roles of the girls in bullying process. Performing these kinds of analyses with sample groups in which the girls are at quorum is another suggestion of the study. The long term success of the bullying proofing programs is possible with the cooperation of teachers, school staff, family members and children.

References

Bademci, V. (2006). Güvenirliği doğru anlamak ve bazı klişeleri yıkmak: bilinenlerin aksine, cronbach'ın alfa katsayısı, negatif ve "-1"den küçük olabilir. *İnönü Üniversitesi Eğitim Fakültesi Dergisi, 7*(12), 3-26.

Bauman, S., & Del Rio, A. (2005). Knowledge and beliefs about bullying in schools: comparing pre-service teachers in the United States and The United Kingdom. *School Psychology International*, *26*(4), 428-442.

Beran, N.T. & Tutty, L. (2002). An Evaluation of the dare to care: bully proofing your school program, final report. Retrived from http://www. ucalgary.ca/resolve/files/dare_to_care.pdf.

Byrne, B. (1994). Bullies and victims in a school setting with reference to some Dublin schools. *The Irish Journal of Psychology*, *15*(4), 574-586.

Connor, D.F. (1998). Over categorical aggression in referred children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, *67*(1), 66-73.

Çınkır, Ş. & Karaman Kepenekci, Y. (2003). Öğrenciler arası zorbalık. *Kuram ve Uygulamada Eğitim Yönetimi*. *34*, 236-253.

Dash, M. (2007). *Education of exceptional children*. Atlantic Publishers and Distributors (p) LTD, New Delhi.

Düzgün, E., Alibeyoğlu, M.C. & Orhan, F. (2006). "Şiddete karşı farkındalık" öğretim programının etkililiği ve öğrenci görüşleri. Bildiri Özeti. I. Şiddet ve Okul: Okul ve Çevresinde Çocuğa Yönelik Şiddet ve Alınabilecek Tedbirler Sempozyumu, 28 - 31 Mart 2006, İstanbul.

Ersanlı, E. (2007). *Okullarda şiddet ve çocuk suçluluğu*, Adem Solak (Ed). Ankara: Anı Yayıncılık.

Eslea, M. & Rees, J. (2001). What age are children most likely to be bullied at school? *Agressive Behavior*, *27*(6), 419-429.

Garrity, C., Jens, K., Porter, W., Sager, N. & Short-Camilli, C. et al. (2000). *Bully-proofing your elementary school*, (2d ed.) . Longmont, Colorado: Sopris West.

Griffin, R.S. & Gross, A.M. (2004). Childhood bullying: Current findings and future directions for research. *Aggression and Violent Behavior*, *9*, 379-400.

Kartal, H & Bilgin A. (2007). İlköğretim öğrencilerine yönelik bir zorbalık karşıtı program uygulaması: okulu zorbalıktan arındırma programı. *Eğitimde Kuram ve Uygulama. (Journal of Theory and Practice in Education). 3*(2), 207-227.

Kartal, H & Bilgin A. (2008). Öğrenci, veli ve öğretmen gözüyle ilköğretim okullarında yaşanan zorbalık. *İlköğretim Online*. *7*(2), 485-495. Retrived from http://ilkogretim-online.org.tr.

Kartal, H. (2009). öğretmen adaylarının uygulama okullarındaki zorbalıkla ilgili değerlendirmeleri. *Gazi Eğitim Fakültesi Dergisi*. *29*(1). 141-172.

Kırcaali-İftar, G. (2007). *Özel gereksinimi olan diğer öğrenciler.* S. Eripek (Ed.), İlköğretimde Kaynaştırma. Eskişehir: Anadolu Üniversitesi Yayınları, 125-137.

Moore, K.A. (2006). Defining the term at risk. *Researchs to Results Child Trends.* http://www.childtrends.org/Files/DefiningAtRisk%5B1%5D.pdf.

Olweus, D. (1987). Schoolyard bullying - Grounds for intervention. *School Safety*, Fall:4-11.

Olweus, D. (1993). *Bullying at school* – What we know and what we can do. Oxford: Blackwell Publishers.

Olweus, D. (1999a). Sweden. In Smith, P.K., Morita, Y., Junger-Tas, J., Olweus, D., Catalano, R., & Slee, P. (Eds). *The Nature of School Bullying: A Cross-National Perspective*. London & New York: Routledge, 7–27.

Özkan, Y. & Gökçearslan E. (2010). Düşük sosyo-ekonomik düzeydeki ilköğretim okullarında akran zorbalığı. *İlköğretim Online*. *9*(2), 576–586. Retrived from http://ilkogretim-online.org.tr.

Pepler, D.J., Craig, W.M., Ziegler, S. & Charach, A. (1994). An evaluation of an anti-bullying intervention in Toronto schools. *Canadian Journal of Community Mental Health*, *132*(2), 95-110.

Sarpdağ, M. (2005). *Çocuk suçluluğu ve polis*. Ankara:Ahsen Matbaacılık.

Scarpaci, R. (2006). Bullying: Effective strategies for its prevention. *Kappa Delta Pi Record*, *42*, 170-174.

Schatz, J.N; Smith, L.E; Borkowski, J.G; Whitmen, T.L. & Keogh, D.A. (2008), Maltreatment risk, self-regulation and maladjustment in at-risk children. *Child Abuse And Neglect*. *32,* 972-982.

Smith, T. E., Polloway, E.A., Patton, J.R., & Dowdy, C.A. (2004). *Teaching students with special needs in inclusive settings* (4th ed.). Boston, MA: Pearson.

Stevens, V., de Bourdeaudhuij, I, & Van Oost, P. (2000). Bullying in Flemish schools: an evaluation of anti-bullying interventions in primary and secondary schools. *British Journal of Educational Psychology*, *70*, 195–210.

Uysal, A. (2006). Şiddet karşıtı programlı eğitimin öğrencilerin çatışma çözüm, şiddete eğilim ve şiddet davranışlarına yansıması. Bildiri Özeti. I. Şiddet ve Okul: Okul ve Çevresinde Çocuğa Yönelik Şiddet ve Alınabilecek Tedbirler Sempozyumu, 28 - 31 Mart 2006, İstanbul.

This page is intentionally left blank

iejee | INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

Learning Experiences and Practices of Elementary Teacher Candidates on the Use of Emerging Technology: A Grounded Theory Approach

EunJin Bahng[a], Mimi Lee[b,*]

Received: 14 August 2017
Revised: 06 Sept. 2017
Accepted: 18 Sept. 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236118

Abstract

The purpose of this study is to understand the phenomenon of the "professional journey" of elementary teacher candidates (ETC) both as learners and as teachers by exploring their learning experiences and practices regarding the virtual reality (VR) platform called Second Life (SL). Using the grounded theory approach, we designed an insider-led work-based study, synthesizing an outsider perspective. We collected data from 229 ETCs enrolled in science methods course over three academic years including semi-open ended reflective questionnaire, science journal notebooks, 5E science lesson plans, peer teaching classroom observations, and instructor's semester notebooks. Results indicate that as a VR platform, SL provided unique learning experiences that encourage discovering, problem solving, or accidental learning for these mostly first-time user ETCs. Three themes emerged in terms of perceptions of integrating VR into science teaching: (1) skeptical integrators (29%), (2) observant integrators (59%), and (3) innovative integrators (12%). However, none of the ETC collaborative groups integrated a VR platform during their 5E science lesson planning activities, demonstrating the gap between learning experiences and practices. Finally, we converged analysis results with related literature to develop a theory. This study contributes to the discussions on preparing teachers for the 21st century in which all will live in a technology-accelerated society.

Keywords: Elementary teacher candidates, technology integration, emerging technology, virtual reality, second life.

Introduction

Digital citizenship is no longer optional, but necessary, as we are living in a fast-evolving era of the Information and Communication Technology (ICT) world. This revolutionary movement from an analogue-based to a digital-based society has also changed the types of occupations needed; the roles of creative individuals as the key agents of change; the ways people collaborate with others and interact with digital devices; and, the ways ideas are developed, shared, and distributed (deSessa, 2000; Friedman, 2005). For some scholars, our digital-based society is recognized as an *exciting time* and as a *powerful catalyst* for rethinking innovative ways to teach and learn science (deSessa, 2000; Slotta & Linn, 2009).

For many decades, the needs and demands for preparing the workforce in the fields of science, technology, engineering and mathematics (STEM) have increased; at least this is the case in the United States (Kuenzi, 2008; National Commission on Excellence in Education [NCEE], 1983; National Academy of Science [NAS], 2010). The national efforts to improve student achievement in science and mathematics have also intensified in the U.S. (National Science Board [NSB], 2006). For many years, teacher shortage and inadequate preparation of highly qualified teachers in these areas have remained a challenge for the nation (The California Council on Science and Technology [CCST], 2007; Triangle Coalition for STEM education, 2017). Still, there is a limited collective understanding on how to increase the quality of STEM education, how to prepare highly qualified science and math teachers, and how to enlarge the STEM workforce.

The role of teachers is considered a critical component in STEM fields (CCST, 2007). Some believe that high-quality teachers would induce meaningful learning and, subsequently, increase the number of students who choose STEM career paths (NSB, 2006). This, in turn, would partially solve the challenge of high demand for STEM-skilled workers in the U.S. (NSB, 2006; U.S. Congress Joint Economic Committee, 2012). In spite of this well documented need, the CCST report acknowledged the fact that there is a lack of well-prepared and effective teachers who teach science and mathematics at the secondary level. Furthermore, less than one percent of the available federal STEM resources for teacher performance improvements have been dedicated to preparing future teachers—only $0.43 million out of $312 million available (National Science and Technology Council [NSTC], 2011). Among these investments, there were only five programs that focused on STEM fields for K-8 pre-service elementary teachers (NSTC, 2011).

Some teacher education programs (TEP), however, shed light on preparing prospective teachers. For instance, Klenier, Thomas, Lewis, and Greene's (2007) report indicates that more than 90 percent of the identified TEP

[a,*]Corresponding author: EunJin Bahng, School of Education, College of Human Sciences, Iowa State University, 1740D Lagomarcino Hall, 901 Stange Road, Ames, Iowa 50011-1041, USA. E-mail: ejbang@iastate.edu
[b]Department of History Education, Hongik University, 94 Wausan-ro, Mapo-gu, 04066, Seoul, South Korea. E-mail: mimilee@hongik.ac.kr

were already integrating technology into instruction, which ranged from utilizing Internet resources to content specific software tools. 93 percent of the participating institutions reported that they teach educational technology within methods courses. When investigating factors that were considered as barriers to integrating educational technology into programs, pre-service teachers' lack of interest and limited skills and knowledge were not considered as significant barriers during program coursework or field experiences. When it comes to integrating technology into field experiences, competing priorities in the classroom and available technology infrastructure in the schools were the two most frequent barriers to the participating institutions (Klenier et al., 2007).

The Purpose of Study

The elementary teacher candidates (ETCs) in this study were in a teacher preparation program where digital citizenship and technology integration with instruction were highly valued. Towards the end of the 4-year program, the ETCs take a science methods course that was designed for them to engage in a new kind of learning and instruction using emerging technologies, specifically a Virtual Reality (VR) platform within which the ETCs are avatars. This study explored the journey of these ETCs, specifically their learning experiences and practices regarding a VR platform called Second Life (SL), which is a 3D digital world. Adopting the grounded theory approach as a research method, this study was structured as an insider-led work based study while synthesizing an outsider perspective. The questions central to this study were:

1. What was the nature of the learning experiences when elementary teacher candidates were engaged in VR activity?

2. How did elementary teacher candidates perceive integrating VR platforms into their science teaching?

3. What did the teaching practices of elementary teacher candidates look like when they integrated VR platforms within the 5E instructional model?

Working Definition of Elementary Teacher Candidates (ETCs).

ETCs exhibit many characteristics and these are well identified in the literature (e.g., Bleicher, 2006; Crowther & Bonnstetter, 1997; Newman, Abell, Hubbard, McDonald, Otaala, & Martini, 2004). ETCs enter the science methods course with formal preparation from only two or three courses in the life sciences and physical sciences during secondary school or university studies; therefore, it is likely they lack sufficient subject matter background. They may have negative and/or indifferent attitudes towards science and science teaching and, as a result, may lack confidence in science teaching. The students often are not well-versed in inquiry-based science teaching processes that promote the use of evidence, logic, creativity, reasoning, thinking, and communication skills, possibly because they are heavily focused on their grades and many would prefer to be taught in a didactic way (i.e., simply memorizing facts and taking tests) that is more comfortable to them. During the course, many may feel they have learned nothing from inquiry-based learning and will struggle to identify concepts they have learned through inquiry-based learning. Finally, most believe that there are always correct answers in science (Newman et al., 2004).

Methodology

The Grounded Theory Approach

This study adopts the grounded theory approach. In 1967, Glaser and Strauss established systematic and scientific guides in qualitative methods called "grounded theory" (GT) in contrast to quantitative methods (Dunne, 2011). According to Glaser and Strauss (1967), GT is a research method influenced by symbolic interactionism for developing a theory that conceptualizes the specific social concepts, patterns, and structures through the process of constant comparative methods. GT is interested in knowledge or reality founded in empirical data (Johnson & Christenson, 2012).

This study is based on Glaser's (1998) grounded theory in that the literature review—existing theories of the field—was done after collecting data, and integrated into discovered categories from the empirical data to develop and explain properties of the discovered categories and variables. Glaser (1998) claims that these discovered core categories are what legitimize GT work since these core variables open up existing theories and suggest formal and creative aspects of the core variables for future research.

As guided by Gasson's (2004) reflexive grounded theory model, we acknowledge the influence of literature sources in order to build the contextual needs for the existence of the current study as well as our own prior knowledge of the topics. We also recognize the influence of literature sources in developing the conceptual framework in the study. Yet, the literature review related to the main questions was considered as secondary data and was integrated after the core themes of the study were discovered (Gasson, 2004). As a result, the core themes from the first and the secondary data sets are used for the formal theory construction and final interpretation. Table 1 summarizes how three research foci—the nature of *learning experiences* in an emerging technology activity, perception of *integrating emerging technology* into science teaching, and *teaching practices* when integrating emerging technology with science—were aligned with the data sources.

Context of The Study

The two researchers and their membership roles. The first author taught the science methods course as the sole instructor. This study, thus, qualifies as an insider-led work-based study (Costley, Elliot & Gibbs, 2010). The researcher (1) had certain degrees of control in initiating changes, direction of discussions, and critical inquiry as the main facilitator; (2) was able to practice critical thinking and self-reflection during the investigation; and finally (3) involved elements of personal and professional development for all the participants in the study (Costley, Elliot & Gibbs, 2010; Reed & Proctor, 1995). Yet, the researcher's role during the VR activity was limited to that of a peripheral member researcher (Adler & Adler, 1987; Dwyer & Buckle, 2009).

The first author taught the science methods course more than five years. She engaged in continuous curriculum developmental efforts based on her own reflections, outside feedback, and current reform efforts. As a result, revisions and adjustments in the curriculum were made each semester. In addition, the first author engaged in weekly VR activities for two years within the SL platform with a colleague in literature education, optimizing learning experience through a virtual property in the New Media Consortium (NMC) in SL (e.g., CyberTechs, a location for the island: Teaching 2, 179.96.23).

The second author, as an outside member, conducted the literature review alongside the investigation. As noted in grounded theory literature, it is impossible to know beforehand what will emerge from the investigation. Scholars also caution researchers not to be confined by previous theories (Corbin & Strauss, 2008; Dunne, 2011; Glaser & Strauss, 1967). As the investigation evolved and relevant concepts emerged from the data, the literature review in this study revealed different aspects. The literature review, as a result, sometimes confirmed findings, but at other times illustrated discrepancies between our findings and existing literature.

Table 1. *Summary of the design of the study*

Grounded Theory Approach						
		An insider-led work based study while synergizing outside perspectives (emic and etic insights)				
Major goal	Questions	Study Context	Study Participants	Data Collection	Constant Comparative Method	
Understand phenomena of professional journey of elementary teacher candidates by developing a theory on the learning experiences and practices in regards to emerging technology	1. What was the nature of learning experiences when elementary teacher candidates were engaged in VR activity? 2. How did elementary teacher candidates perceive integrating VR platforms into their science teaching? 3. What did the teaching practices of elementary teacher candidates look like, if they integrated VR platforms with science?	Public university North central region Senior-level Science methods exit course Three academic years	229 elementary teacher candidates (1st year 70, 2nd year 97, 3rd year 62) Female: 205 Male: 24 European-American: 224; African-American: 3;Hispanic Origin: 2 Elementary education major: 221 Early Childhood education major: 8	Semi-open ended reflective questions (Total of 227) Science journal notebooks (Total of 217) 5E science lesson plans (Total of 47) Peer teaching observations (Approx. 42 hours) Instructor's semester notebooks (Total of 7) Secondary data Literature review	Analytic Induction Comparing incidents applicable to each category/variable ⇩ Integrating categories and their properties ⇩ Delimiting the theory (parsimony of and theoretically saturated core categories) ⇩ Discovery of theory	Open Coding Memoing Selective Coding Memoing Theoretical coding/innovation

In the process of forming theories, the first author's insider perspective was augmented by the second author's outside point of view, taking advantage of the different memberships brought to the research (Dwyer & Buckly, 2009). Our two perspectives were reconciled in forming final theories and interpretations.

The Participants

The participants in the study were all enrolled in a teacher education preparation program in a north central region university in the United States. 229 elementary teacher candidates (ETC) participated in the study over three academic years. There were 24 male and 205 female ETCs involved, the majority coming from a European-American ethnical background (98%) and majoring in Elementary Education (97%). Demographic data for the participants are summarized in Table 1.

Science methods course

This is a required exit course for senior level undergraduate students who are enrolled in this elementary and early childhood teacher education program. The course aims to develop ETCs' personal teaching philosophies regarding how students best learn science according to current educational research. The course also emphasizes a practice and reflective approach in order to a) develop a community of active learners, b) learn how to design student-centered and inquiry-based curricula, and c) learn how to evaluate one's own instructional practices.

Disciplinary core ideas include physical sciences, life sciences, and earth and space sciences, along with other concepts such as science as inquiry, scientific and engineering practices, the 5E model of instruction, as well as science, technology, and society—which were informed by *the National Science Education Standards* (NAS, 1996). The course content is also influenced by the *Framework for K-12 Science Education* as well as the *Next Generation Science Standards* (NGSS) (National Research Council [NRC], 2012).

Weekly activities of this course concentrate on the development of ETCs' pedagogical content knowledge. For example, ETCs engage in learning about assessment, learning environments, curriculum, context, classroom management, the nature of science, and socio-culturalism in the context of teaching science in the K-8 grade levels (Veal & MaKinster, 1999). Further details on the specifics of the weekly activities of the course can be found in Bang (2013a).

Conceptual Model for Purposeful Activities

ETCs in this study participated in three purposeful activities: VR activity (Quest: We Are SciTeachers!), 5E

science lesson plan activity, and peer teaching activity. The activities were designed within the frameworks borrowed from Wenger's (1998) Communities of Practice (CoP) and Wells's (2000) dialogic inquiry. Guided by these two frameworks, the activities were then engineered to routinely model Engestrom's (1987) cultural-historical activity theory and Freire's (2003) action and reflection cycle. This conceptual model assumes that the essence of learning is practicing and participating in purposeful activities within the community life and apprenticeship as opposed to the authoritarian tradition of lecturing (Freire, 2003; Holland, Lachicotte, Skinner, & Cain, 1998; Lave & Wenger, 1991). The conceptual model also assumes that it is a systematic way to *humanize* learning and teaching (Freire, 2003).

Learning within the concept of CoP involves recognizing the existence of a community as a shared enterprise and having a set of culture-specific practices. The members in this community actively produce new meaning through negotiations and practices. Successful members, non-marginalized ones, ultimately become core members of the community while shifting knowledge, skills, roles, motives, cultural capital and tool-kits, values, and identities in practice (Bourdieu, 1986; Wenger, 1998). In addition, learning within the CoP rejects the idea of learning as pre-planned, but rather as something that "belongs to the realm of experience and practice" (Wells, 2000; Wenger, 1998, p. 225).

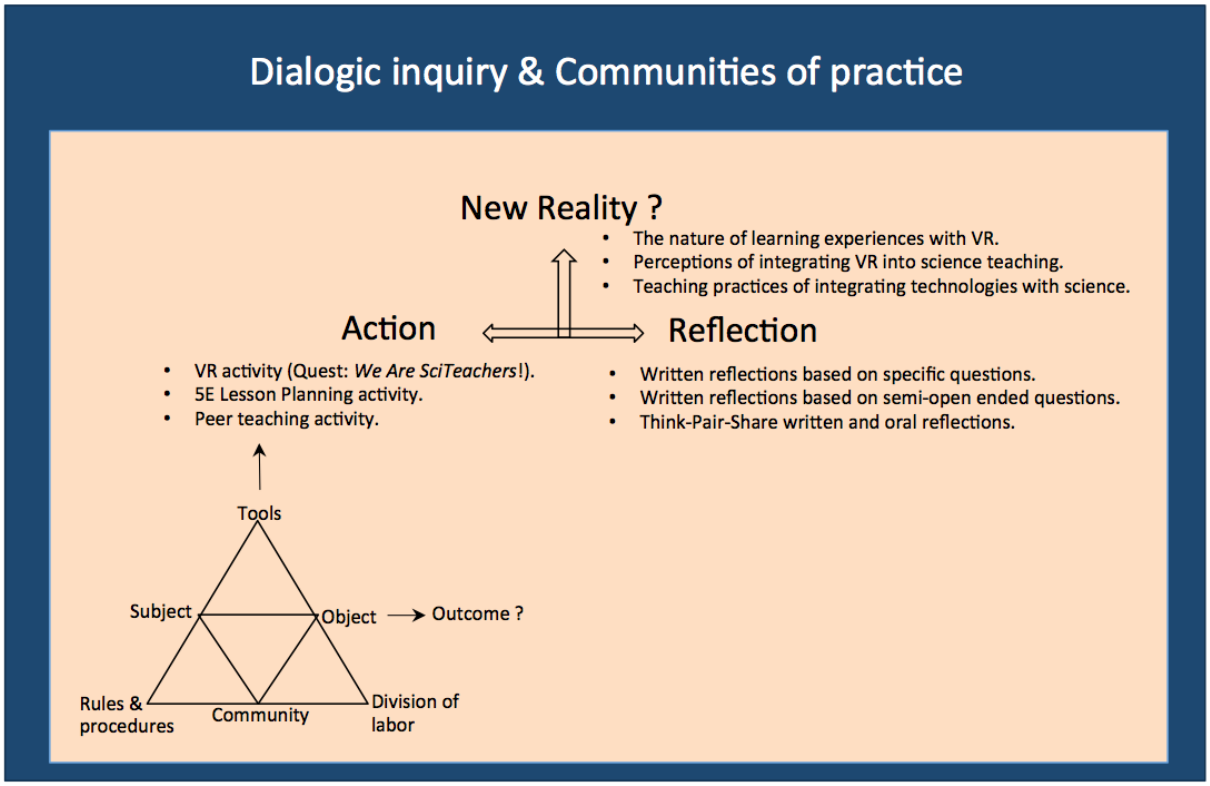


Figure 1. Conceptual model of the purposeful activities situated within the multiple learning theories

The concept of dialogic inquiry, aligned with sociocultural theories, embraces the constructive nature of learning and the learning-by-doing approach that emphasizes intrapersonal-interactive collaborative learning, meaningful experiences, and mentors who follow a scaffolding process (Applefield, Huber, & Moallem, 2001; Dewey, 1997; Greenfield, 1999; Rogoff, 1990; Vygotsky, 1978). According to Wells (2000), knowledge is built through co-generative discourse. Bakhtin (1986) considers all discourse as dialogic in that it involves generating common understanding, borrowing others' utterances, containing dialogic overtones, and building collective knowledge and constantly producing meaning through various modes of dialogue. Learning in dialogic inquiry, as opposed to didactic instructional methods, regards learning as process, activities as formation of identity, curriculum as a means to engage in activities, and a classroom as a collaborative community (Wells, 2000). Also, learning in dialogic inquiry demands individuals and communities be open to diversity and originality in solving new problems (Wells, 2000).

During the methods course, the ETCs were guided to explicitly and routinely model Engeström's (2000) cultural-historical activity theory (CHAT) as well as Freire's (2003) action-reflection empowerment cycle. CHAT is originated from Vygotsky's (1978) activity theory and consists of seven elements: subject, object, outcome, instruments/tools, rules, community, and division of labor. According to Engeström (2000), the CHAT framework converges learning organization, knowledge management and social capital. Learning in CHAT means that activities are not only goal-directed but also object-oriented. Therefore, learning activities have coherency and continuity (Engeström, 2000). Yet, these aspects of CHAT generate disturbances and contradictions when deviations from the original scripts occur. Due to these two rather opposing features of

CHAT, learning in a CHAT context encourages collaborative inquiry, thereby creating a shared vision and collective solutions (Engeström, 2000).

Lastly, Freire's (2003) action-reflection model is adopted to empower ETC to author new realities as a healthy community member and as a self-regulated and critical thinker (Freire, 2003; Wallerstein, 1998). This model promotes knowledge being transformed through lived experiences in order to unveil new knowledge. Learning in the action-reflection cycle promotes development of experiences, critical discovery, openness, and de-mythologizing knowledge. This model also puts dialogue at the very center of learning and knowing to critically confront reality, and to establish a horizontal relationship among community members (Freire, 2003).

The three purposeful activities: Virtual reality, 5E science lesson plan, and peer teaching activities

The ETCs participating in this study engaged in three purposeful activities related to VR and science teaching over the course of a semester. The common goals which aligned with the conceptual model (see Figure 1) for the three purposeful activities were (a) to develop ETC's capacity as a self-regulated learner and critical thinker through various reflective practices such as self-assessment, self-reflection, and interactions with critical friends within multiple communities of practice; (b) to learn effective methods of science instruction and assessment; and finally, (c) to build a knowledge base about science teaching.

Prior to the three activities, the ETCs were engaged in modeling both the conceptual model (see Figure 1) and the 5E learning cycle (Bybee et al., 2006). Specifically, they were given multiple learning opportunities to practice how to be an active community member; how to define each other's roles, rules, and procedures in a collaborative learning strategy; and, how to design a 5E science lesson plan using disciplinary core ideas. The ETCs participated in the three activities sequentially over six weeks. ICT tools used during the activities were the following: emails, Blackboard Learn, PBwiki (my.pbworks.com), and Second Life (www.secondlife.com).

Virtual reality activity: QUEST; We Are SciTeachers! This VR activity is designed to develop 21st century skills towards technology, including elements of playfulness, collaborative teamwork, judgment, multitasking, networking, negotiation, a spirit of exploration, and imagination (Jenkins, Clinton, Purushotma, Robinson, & Weigel, 2006). A week before participating in the quest, the ETCs were informed about the main technology platform that was used, called *Second Life* (SL, www.secondlife.com), a currently available free massively multiplayer online role-playing game (MMORP) program. As a pre-VR activity, the ETCs were individually and along with guidance from the instructor expected to learn the basic features of SL such as safety guidelines, how to create an avatar, and the core rules of netiquette. The instructor first demonstrated her own avatar within the SL platform in terms of how to walk, how to chat, and how to navigate science-related places. Each student explored the SL platform as an avatar once they chose one from the dummy avatars list that contained IDs and passwords. These dummy avatars (female avatars 25, male avatars 5) were provided by the instructor and re-used each semester. Right before the VR activity, the ETCs were given basic guidelines, the science-related places in SL, and the quest sheet, which includes reflection as a post-activity (See Appendix). An alternative activity was also available for those who chose not to participate in the VR activity.

5E science lesson plan and peer teaching activity

These two activities are designed to provide opportunities for ETCs to *design* and *enact* the 5E model of instruction collaboratively (Bang, 2013b). As for the pre-activities, the ETCs engaged in the facilitator-modeling inquiry activities to learn how to design and teach science using a 5E model. These modeling activities were followed by weekly science journal reflections. The ETCs also read selected articles about how to use the 5E model, engaged in a video case study, and designed a mini 5E lesson. The two tools used during the activities were the *Reformed Teaching Observation Protocol (RTOP)* (Piburn et al., 2000) and *Science Teacher Learning from Lesson Analysis (STeLLA)* (Biological Sciences Curriculum Study [BSCS], 2011).

During the lesson plan and classroom enactment period, the ETCs were given 5E lesson plan guidelines that contained ten items: Introduction, Standards, Objectives, Scientific background, Common misconceptions, Diversity and inclusion strategy, Classroom management, Materials and equipment, Instructional strategy, and a 500-word Individual reflection. The 5E science lesson plan activity is an evolving and iterative process in that the ETCs, as a team (a group of three or four), participated in outlining, writing a 1st draft, a 2nd draft, and a final draft of their 5E science lesson plan before their peer teaching. Integrating technology was not a required item for their 5E lesson planning and teaching but was encouraged through verbal and written feedback. During the peer teaching activity, a group of peer panel members and the first author made observational notes based on RTOP and STeLLA tools while a peer teaching group and the class were engaged in science teaching activities. Peer teaching typically lasted for 50-60 minutes. Right after the peer teaching activities, the whole class was engaged in an in-depth discussion regarding the lesson plan and its implementation, science content taught, classroom culture, 5E sequence, teacher questionings, and/or science content storyline. Finally, each team received a written feedback summary that reflected the comments from the panel members, the first author, and the whole group discussion.

Data Collection

In this study we collected and analyzed data from five types of data sources: semi-open ended reflective questions, science journal notebooks, 5E science lesson plans, peer teaching classroom observations, and instructor's semester notebooks. The data were collected after ETCs received their final grades. To maintain objectivity and anonymity, three research assistants digitized the raw data and assigned numbers from 1 to 229 for the archived data before the first author and a research assistant began the analysis.

Both verbally and in writing, the participants were guided to reflect upon their learning experiences right after the VR activity. Four semi-open questions were provided for individual reflection. The participants uploaded these reflections to the class wiki page. Each week, the participants were guided to make half-page journal entries in a composition notebook related to the topics discussed. There were 11 weekly journal entries per semester. A total of 217 of 229 science notebooks were collected, assigned random numbers, scanned, and cataloged in a Microsoft Excel spreadsheet. Twelve science notebooks were removed from our data set due to handwriting illegibility issues or having more than three missing journal entries. As for the data from the VR activity, twenty-two female participants chose not to participate in the activity due to physical dizziness or other personal reasons.

As for the 5E science lesson plans, only the final drafts were collected (a total of 47 science lesson plans). These science lesson plans were scanned and cataloged in a Microsoft Excel spreadsheet recording targeted events and categories (e.g., Was a technology used or not? If so, why and how was it used? What technology was used and at what stage? Was it used in the 5E learning cycle?). Approximately 42 hours worth of peer teaching observations were made during the science methods courses by the first author. The first author took field notes while observing peer teaching in the semester notebook describing the ETCs enactments, important components, and questions that arose during the observations (a total of seven notebooks). Finally, these instructor's semester notebooks were used to guide the follow up discussions along the lines of Piburn et al.'s (2000) RTOP and BSCS's (2011) STeLLA tool.

Data Analysis

The data analysis involved a constant comparative method under the inductive reasoning process. This study followed four steps that were established by Glaser and Strauss (1967) and Gasson's (2004) reflexive GT approach. These systematic analytical steps consisted of (a) open coding for comparing incidents applicable to each category and variable, (b) integrating categories and their properties, (c) selective coding in order to delimit the theory to abide by the rule of parsimony and theoretical saturations of core categories, and finally (d) theoretical coding and a write-up. One of the critical activities used during the process was memoing, which is viewed as a central element in the GT approach (Dunne, 2011; Glaser, 1998). Microsoft Excel spreadsheets and post-it notes were used to color-code and create memos. Finally, the existing relevant literature review was considered as our secondary data and was integrated into findings when examining for the final theoretical construction.

Trustworthiness

Guided by Shenton's (2004) and Guba's (1981) strategies, this study used multiple data resources in many forms. As a result this study represented multiple voices, had a variety of informants, and therefore reduced the investigator bias (Miles & Huberman, 1994). For instance, reflective questions and science journal notebook data were individual-based, whereas 5E science lesson plans and peer teaching observation data were group-based. The instructor's semester notebook data added additional perspective in understanding the phenomena under inquiry. Additionally, data were collected over three academic years; thus, this study used longitudinal data.

To ensure honesty, the participants were allowed to choose an alternative activity during the VR activity. There was no extra credit given for integrating VR platforms into their teaching practices, nor were the participants penalized for not integrating VR platforms into their 5E science lesson plans and peer teachings. The participants were also asked to collect their science journal notebooks once their grades were posted, and were sent email reminders about the study before considering their science journal notebooks as data for this current study. Srivastava and Hopwood's (2009) iterative framework states that "patterns, themes, and categories do not emerge on their own. They are driven by what the inquirer wants to know and how the inquirer interprets what the data are telling her or him" (p.77). Finally, the two researchers integrated both insider (emic) and outsider (etic) perspectives in building a formal theory (Morris, Leung, Ames, & Lickel, 1999).

Limitations

Regardless of the efforts made to abide by the rigorousness of the GT approaches, our study was bounded by several limitations in data collection and analysis. For instance, the participants were selected based on purposeful sampling, not random sampling. The participants were not asked to assess the trustworthiness of the data or the results. This absence of member checking may influence the validity of the study. Also, the first researcher's role as an insider and the adopted theoretical lenses may affect the researcher bias in understanding the phenomena. Therefore, what the participants experienced during the study was filtered through the researcher's frames of reference. Finally, there were no systematic measures used to learn about the previous familiarity with the VR platform or with MMRPG, rather the participants reported freely. Due to these limitations, the findings of the study should be understood and applied with discretion.

Results

Q1. What was the nature of the learning experiences when elementary teacher candidates were engaged in VR activity?

Theme 1: First time VR users learned by playing around as avatars. The ETCs in this study reported that they were first time VR users who had not been exposed to any kind of virtual reality. During this new experience with a 3D digital world, the ETCs started their learning journey with great frustration. The intensity of the ETCs' feelings of difficulty in this VR platform were the highest during the first 10 minutes in SL. During this initial period, the ETCs tried to make their avatars move, walk, dance, fly, or talk at SL Orientation Island where they saw avatars visiting from all over the world. This level of frustration was also evident for those who had some previous experience with Sim Games. Two typical remarks are as follows:

> At first, it was very difficult to find my way around and maneuver my robot [avatar]. As I continued exploring Second Life, I was able to find very interesting icons that allowed me to get around a lot easier [VR reflection M-ID19].

> My confidence at first was shattered … really I didn't understand anything. At first it was very frustrating but after I practiced and gained a little patience then it didn't seem so bad. After this I feel a bit more confident with SL. [VR reflection F-ID6]

The instructor's semester notebook indicated the need for modifying pre-VR activities by having more first-hand experiences with a 3D digital platform. Interestingly, these initial feelings of frustration and difficulty rapidly turned into enjoyable and engaging explorations. The ETCs felt that this was due to the fact that the VR platform allowed them to practice and to gain patience and confidence with SL. After playing around with the program for about 20 minutes, the ETCs started exploring and discovering further interesting features.

Theme 2: First time VR users learned by discovery, by problem solving, or by accident. The ETCs predominantly said that they experienced a true form of discovery and problem solving through the VR activity. Initially, the ETCs as avatars considered SL as an unknown world since they had to learn the language, software features, and rules of engagement. The ETCs as avatars had to play around with many features within the SL program such as how to group chat, how to create an e-note card, how to change the appearance of their avatars, and how to read a mini-map. Also, the ETCs as avatars had to observe their visually intensive 3-dimensional surroundings and make collective or individual decisions on what to do and how to navigate their surroundings in SL. Some described their experience as "figuring things out until you get the hang of things." Representative comments are:

> I grew more in my digital skills and confidence by discovering things dealing with Second Life. I still do not feel totally confident using the program because there is so much that goes with using Second Life. [VR reflection F-ID47]

> My avatar visited a world whose theme was based on futuristic settings and equipment. Many of the things to do in Second Life cost money but I was able to drive an ATV type vehicle for free. Another aspect that related to science content was investigating the waters and flying through the trees. The visuals are very realistic so it was neat to swim in ponds and the ocean. There were also flowers and plant life but the avatar was unable to physically interact with these things. When first investigating Second Life I had to use inquiry skills such as trial and error or asking experienced avatars how to complete an action or activity. [VR reflection F-ID219]

Before the activity, some of the frequently used SL specific vocabularies were reviewed (e.g., avatar, islands, teleport, friendship, object, map, main grid, freeze, prim, machinima, Linden dollar, skin, texture, in-world, and off-world). One comment that illustrates ETCs learning vocabulary terms is "I learned the word *teleport* to travel to new islands." Teleport is a way for avatars to instantly travel from one island to another. The ETCs were able to understand and use these terms after exploring, or after a lot of trial and error for themselves. In this way, these hands-on activities encouraged rapid learning of SL terms. Some ETCs considered their SL learning experiences similar to inquiry, the 5E learning cycle, or experimentation. Most ETCs valued the learning by discovery that the VR platform offered and recognized that their digital skills had increased.

Theme 3: First time VR users learned science facts and digital skills by social interactions with 3D content and with other avatars. The ETCs navigated science-related islands both as a group and as individuals. In order to function as a team, the ETCs needed to offer and accept each other as friends. This enabled them to instant message, personal chat, and group chat. They could also send each other a teleport invitation. This function can allow any avatar in the friendship list to move instantly to the inviter's island. During this collaborative exploration, they as avatars asked a lot of questions of each other through a nearby chat, a personal chat or a group chat, and suggested some group actions (e.g., I found a cool island! Let me teleport you, Let's fly together to look at the stars in the sky, Let's run to the inside of a stomach). One illustrative comment was:

> I was the reporter for our group. My job was to create short memos about each place we went to. It was hard at times because while I was typing the notes my group members would find another great place without me. After we found out how to transport each other then this problem was solved. [VR reflection F-ID28]

Once they teleported to a science-related island, the ETCs actively engaged in deciding what to explore and how to explore the island. Some engaged in conversations with other avatars describing what they were doing with a nearby chat. Some ETCs found this new way of social learning challenging due to the fact that there were so many great places to explore within an island and there were other unknown avatars walking around. Still others initially were unsure of what to do once they teleported themselves to a science-related island. Some ETCs initiated conversations with nearby avatars that were not in their friendship-list when they needed help about how to use some features in SL. Two comments illustrating such social interactions are:

> This made me realize that in my three hours of investigating this virtual world, I had only experienced a glimpse, which made the confidence in my virtual reality skill deflate quite a bit. I was, however, proud of the confidence I gained by interacting with other avatars. [VR reflection F-ID32]

> We got space suits and saw the moon. This was the first place we landed so we were learning how to communicate with each other and explore [Apollo Moon Landing Island]. We did not think it was very thrilling, so we went to Cloud Chateau. When we arrived we were very confused. We looked around and learned how to sit in chairs. We fell from the cloud and landed on an island and walked on the sea floor. We learned about using angles. …At Science Friday, we encountered another avatar. We talked to him a little and explored. We saw a penguin but did not figure out how to communicate or do anything with him. We also found that we could teleport by sitting in different chairs. When we teleported to Neil Armstrong Library and Archives, there was a person [an avatar] in the museum to take us to a website. We were able to read articles and learn different historical facts. … We found some other classmates at

> Virtual Stomach and rode on a Ferris wheel. We found green houses full of peas. You can play a fun chemistry 101 Game on this area... [VR group 34 notecard]

While learning and defining group roles and networking, the ETCs naturally learned digital skills. The ETCs explored their selected science-related islands *by looking at different things* such as 3D dinosaur models, a spaceflight museum overlooking planet Earth, habitats, weather maps, rockets, or animals. Secondly, the ECTs also engaged in SL *by interacting with the 3D objects* such as conducting experiments in a laboratory, using a microscope to see cell samples, watching short video clips by clicking a button, asking questions to guides-avatars, or riding on hot air balloons. Finally, the ETCs gathered relevant facts *by traversing between 2-dimentional websites and 3-dimensional SL*. For instance, when a group of the ETCs visited a museum, they were given a button which directly took them to a website link.

Q2. How did elementary teacher candidates perceive integrating VR platforms into their science teaching?

Theme 1: Skeptical integrators (29%). SL was a frustrating experience. I don't see any merit in using VR platforms. Thus, I wouldn't integrate it into my science teaching. The ETCs in this category demonstrated negative attitudes towards integrating VR platforms into their science teaching. Most of them remarked on the challenging experiences they had with the VR activity. The ETCs in this category disclosed that they did not learn anything scientific, or any new science content, or any useful vocabulary from the VR activity. Three illustrative comments are:

> I am unclear about the reasoning why you would incorporate an activity like this in the classroom. [Science Notebook ID8]

> I do not think this type of lesson would be useful for my class. I was easily confused about how to operate Second Life & didn't think it would greatly increase a student's science learning. [VR reflection M-ID5]

> I did not learn any new science content and vocabulary from this exploration. I was not able to learn about the places, animals, or innovations I saw through the experience. [VR reflection F-ID047]

Therefore, they disregarded VR platforms as a possible learning and teaching tool for their science classroom. The reasons they provided were diverse. For instance, some were suspicious that the systemic problems they encountered in SL would make them vulnerable to viruses. Some disliked the technical problems with slowly loading images. Some feared that if VR platforms were offered to young learners, they would not be able to have solid guidance control (e.g., inappropriate content, unknown avatars behind the computers). As a result, they predicted that this SL could pose a potential danger for their classrooms.

Theme 2: Observant integrators (59%). SL was an interesting experience. I see some merit in using VR platforms. Thus, I might integrate VR into my science teaching with some modifications. The ETCs in this category were open to integrating VR platforms into their science teaching only after some modifications are made. Specifically, the ETCs would surely integrate VR platforms once they practiced a lot more and thought of ways to use it for themselves. Three illustrative comments are:

> I think that SL can be used as a resource for a classroom but I would not use it unless I could provide a focused lesson. There are many distractions in this game and my fear would be that they would end up off task. [VR reflection M-ID23]

> I would have to have a lot more practice in order to think of ways I would use SL. It's very hard to understand and get around. It reminds me of the Sims a lot...If I could find a SL for kids then I would use that in my classroom. It would be nice if my class were the only ones in SL. [VR reflection F-ID21]

> There are alternative virtual worlds, such as RiverCity, that are made for science investigation of, such as, pollution and over-population. RiverCity is created to make a student the scientist in the world and come up with real life solutions to plausible problems that could happen to cities in the future. As stated before, I could see the potential (if highly monitored and it were impossible to come into contact with unfamiliar avatars) because students could build communities, start businesses and build houses. I look forward to further investigating appropriate, well-monitored virtual worlds that I can apply in my classroom to create a safe learning environment. [VR reflection F-ID90]

The ETCs in this category also noted the importance of finding VR platforms that are appropriate for young learners before incorporating it. Other suggested modifications were a focused science lesson first and less visually intensive VR platforms (e.g., class-only or grade-level-only VR platforms). They reasoned that these modifications would prevent students from being distracted or going off-task in their classrooms. The ETCs in this group embraced the initial challenges that they had with SL and valued its features such as interacting with others and with 3D digital worlds as resource and education tools.

Theme 3: Innovative integrators (12%). SL was a fun experience. I see great merit in using VR platforms. Thus, I will definitely integrate VR into my science teaching. The ETCs in this category were enthusiastic about integrating VR platforms into their science teaching. Their reflections indicated much excitement on exploring SL and a desire to try new things along with VR technology, features, and designs. The ETCs considered VR platforms as great mediating tools for young learners to explore on their own, similar to the way they learned it by discovery and social interactions. In integrating VR platforms, the ETCs imagined themselves allowing their students to navigate to new places that were interesting to them and conduct their own learning. Some also perceived that VR platforms would be a great way to show their students the spontaneity and fun in exploration. Three illustrative comments are:

> Once we got into the virtual reality tour, we were able to find many virtual rooms and buildings such as the aquarium and the dinosaur center. It was really fun observing all the areas and becoming more knowledgeable on these subjects. This would be a great way for students to observe and experiment with these

> science related subjects, or also to introduce/become familiar with 5E. [VR reflection M-ID12]
>
> This may make science more interesting for children as they would be allowed to play games. They could possibly learn some information in an interesting way if they stopped to read the posted signs throughout the game...[VR reflection M-ID14]
>
> While using Second Life I learned how to control and navigate my avatar. There are many possibilities of places to travel. There are videos and simulations that you can show your class. If used effectively it could be beneficial to help teach students a topic. [Group8 note card]

The ETCs indicated that VR platforms would be a motivational tool to make science more interesting. For instance, students first could learn how to control and navigate their avatars, and then play games while traveling to and taking photos of many different places. As a teacher integrating VR platforms, they illustrated that they would show videos and simulations to their students when learning about science topics. They would also search for helpful teacher resources within VR platforms that would lead them to 2D websites. Within the 5E model of instruction, some indicated that they would use VR platforms only after their students had acquired first-hand experience and explicitly learned about the scientific concepts.

Q3. What did the teaching practices of elementary teacher candidates look like when they integrated VR platforms with science?

The ETCs seldom integrated technology with science during their teaching practices and rarely mentioned it in their science journal notebooks (only 4% of the notebook entries). When prompted on technology integration, the ETCs mentioned computers, Internet, websites, ELMO, and videos as their technology tools for science teaching. Furthermore, none of the ETCs in the study integrated VR platforms into their 5E science lesson planning or peer teaching, as illustrated in these excerpts from two lesson plans:

> Explore: The students will research about tadpoles and complete the "FACT" sheet (Including: habitat, what they eat, lifecycle, etc.) [5E science lesson plan ID8: The lifecycle of a tadpole, 3rd grade]
>
> During the engagement stage, Teacher will introduce topic: "Today we are going to talk a little bit about volcanoes. What do we know about volcanoes? Has anyone seen a volcano erupt before? We are going to take a look." Teacher pulls up YouTube videos listed under materials. The overall purpose of these videos is to get students interested in finding out why volcanic eruptions differ in power and lava flow speed. Students watch and observe them, writing down observations and inferences in a chart format. [5E science lesson plan ID33: Volcanoes, 5th grade]

In addition to a complete absence of VR integration, only 29% of the 5E teams integrated some type of technology, specifically websites, short video clips, or computers. For instance, the ETCs used teacher pre-screened websites for student research activities mostly during the exploration or elaboration stage. Short video clips were also used for motivating students during the engagement stage or to define major terms during the explanation stage (e.g., phases of the Moon, volcanoes).

Convergence of Results with Related Literature

We conducted the literature review, in part, to further understand our findings from the analysis and, in part, to validate and refine what we found (Corbin & Strauss, 2008; Dunne, 2011). The findings from the analysis revealed the journey of ETCs from a learner of science in a VR activity to a science teacher using VR. As a result, we focused our literature review in three areas, which helped us interpret the analysis results and strengthen our understanding of ETCs' journey from learners to teachers of emerging technologies. The three focus areas included: (a) learning science through emerging technologies in elementary classrooms, (b) learning barriers, experiences, and challenges in integrating technologies into instruction, and (c) preparing elementary science teachers to integrate emerging technologies. Our study specifically focused on ETCs' use of Second Life, which is one of many emerging technologies. When searching for relevant literature, we did not limit the literature to Second Life, in part because we wanted to interpret the findings against a broader backdrop and in part because there is limited research on Second Life in an elementary setting. Figure 2 illustrates the convergence of the major findings of this study with those in the related literature.

Learning science through emerging technologies in elementary classrooms

We were first interested in the use of emerging technologies in elementary science classrooms. When ETCs go out to classrooms, what kind of classrooms would they face? To what extent do teachers use emerging technologies in elementary science classrooms? Reviewing technology use in instruction, Lawless and Pellegrino (2007) stated, "The resources and environments in which teachers and students were expected to use technology prior to 1999 are no longer analogous to current standards" (p. 584). There is, indeed, a wide variety of new and innovative technologies integrated into elementary classrooms nowadays. Video games, augmented reality, and 3D virtual worlds are representative examples of emerging technologies introduced in elementary classrooms.

Several studies reported positive gains through educational video games (Burguillo, 2010; Liu & Chu, 2010). In elementary classrooms, Annetta, Mangrum, Holmes, Collazo, and Cheng (2009) investigated fifth graders' learning of simple machines using a gaming application. They identified gender differences in the time spent for computer and video games. They also reported high student engagement and learning achievements. In-service elementary teachers in a hybrid-mentoring program also reported that emerging technologies, such as various science applications using handheld digital devices, were guaranteed motivational learning tools for their students (Bang, 2013b). Sung and Hwang (2013), focusing on collaborative learning, facilitated Taiwanese sixth graders' sharing and organizing what they had learned with a game-based learning environment called Mindtool. Participating sixth graders improved not only in

attitude and motivation, but also in learning achievements and self-efficacy.

Scholars also identified augmented reality for its potential to enhance learning experiences (Wu, Lee, Chang, & Liang, 2013). Augmented reality provides interactive experiences in a 3D space, but, unlike 3D virtual words where participants are immersed in the 3D environments, augmented reality sits in between real and virtual worlds. Kerawalla, Luckin, Seljeflot, and Woolard (2006) used augmented reality to teach about the Earth, Sun, and day-night cycle to 10-year-olds in the United Kingdom. They identified the potential for augmented reality in educational settings; however, they found that the use of augmented reality did not improve upon a role-playing lesson. Teachers participating in the study also asked for more flexible and controllable materials. Applying augmented reality in an elementary library, Chen and Tsai (2012) reported an improvement in the students' learning performance. They found no gender differences, as compared to conventional librarian instruction.

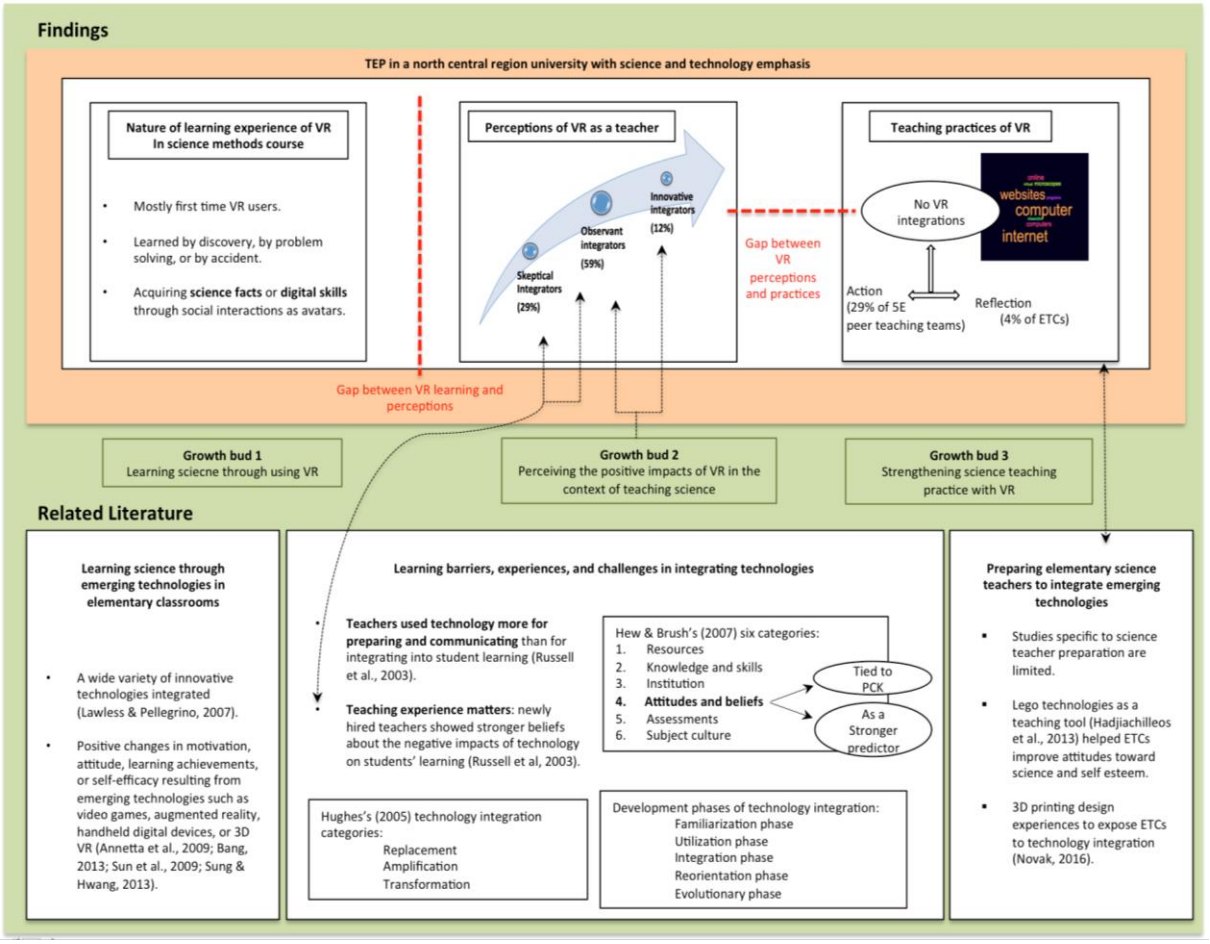


Figure 2. Convergence of the major findings of this study with the related literature.

3D virtual worlds were used to teach plate tectonics to fifth graders (Kim, 2006) and the solar system (Sun, Lin, & Wang, 2010) to Taiwanese fourth graders. Both studies reported higher learning achievements in the 3D groups compared to the control groups. Kim (2006) noted no gender or ethnicity differences in either achievement tests or attitude. SL is often categorized as a 3D virtual world. We, however, had a hard time locating a study using SL in elementary science instruction. Reviewing the research using SL in an educational setting, Inman, Wright, and Hartman (2011) noted that the majority of studies took place in higher education.

Based on the review of emerging technology use in elementary classrooms, we argue that although the research evidence is not sufficient for effectiveness in academic achievement, it is safe to say that various emerging technologies are integrated into elementary science classrooms. The potential of 3D virtual worlds, with their open-ended environments, interactive social tools, and 3D space, also seems promising.

Learning barriers, experiences, and challenges in integrating technologies

When considering emerging technology for instruction, what hinders teachers from using it? Reviewing 48 studies on K-12 technology integration, Hew and Brush (2007) identified a total of 123 barriers that hinder technology integration into instruction and classified them into six categories: (a) resources, (b) knowledge and skills, (c) institution, (d) attitudes and beliefs, (e) assessment, and (f) subject culture. Of these six, teacher attitudes and beliefs are the most relevant to the findings of this study. Teacher

attitudes and beliefs are known to be one of the major barriers to integrating technologies into classroom instruction (Ertmer, 2005; Hermans, Tondeur, Van Braak, & Valcke, 2008). Teacher attitudes and beliefs shape their goals for technology use (Ertmer, Addison, Lane, Ross, & Woods, 1999), affecting the frequency and degree of technology use in instruction. Hew and Brush (2007) argue that teacher attitudes and beliefs influence technology integration more directly, compared to other elements such as subject culture, assessment, and institution. Teacher attitudes and beliefs towards technology are tied to "a teacher's understanding of pedagogy (i.e. pedagogical content knowledge) and to how these various technologies can facilitate learning and achievement among students, and how to assess the various outcomes of learning in these contexts" (Lawless & Pellegrino, p.596).

Teachers' use of technology is also related to teaching experience. Analyzing surveys from 2,894 teachers in 22 Massachusetts districts, Russell, Bebell, O'Dwayer, and O'Connor (2003) found that teachers used technology more often for preparing and communicating than for engaging students in learning activities. In addition, they found that novice teachers (with less than five years of teaching experience) used technology in instruction less than more experienced teachers, although novice teachers demonstrated higher comfort levels with technology. This study also noted, "Teacher beliefs about the importance of technology for teaching was the strongest predictor of the frequency with which technology is used for a given purpose" (p. 302). Novice teachers expressed significantly stronger beliefs about the negative impacts of technology on students learning. Novice teachers, in other words, were less confident that technology would have a positive impact on students in instructional environments, though they used technology to prepare instruction. The negative impacts included, "Making students more lazy, decreasing research skills, and decreasing the quality of student writing" (p.305). ETCs in this study quickly mastered how to navigate and learn science in VR platforms; yet they might not have seen the importance of VR platforms for teaching science.

The research on the stages of technology integration also helps us interpret our analysis results. Stages of technology integration have emphasized the differences and changes in teachers' use of technology; several models of technology integration are available (Toledo, 2005). These stages are not specific to emerging technologies; yet, they do provide insights into how teachers might use emerging technologies in designing and implementing classroom instruction. Types of technology use can be categorized as replacement, amplification, and transformation (Hughes, 2005), each serving different instructional purposes. Further development phases of technology integration include (a) familiarization, (b) utilization, (c) integration, (d) reorientation, and (e) evolutionary phase, although wordings differ slightly per study. The ETCs in this study were familiar with how SL worked, and were able to utilize the technology to perform the quest. However, they fell short by failing to integrate technology into their instruction when they could have used technologies to provide a rich context for students' understanding of pertinent concepts, themes, and processes (Moersch, 1995; Toledo, 2005). The ETCs in this study neither fully reached the integration phase nor moved on to the reorientation or evolutionary phases.

Preparing elementary science teachers to integrate emerging technologies

Finally, we reviewed literature on the preparation of elementary science teachers for technology integration. As access to technologies in classrooms increases, greater emphasis has been placed on preparing elementary teachers to use technology for science education (NAS, 2010). Various methods have been employed either to increase pre-service teachers' knowledge and skills in emerging technologies or to improve pre-service teachers' attitudes towards the use of technologies in instruction. Hadjiachilleos, Avraamidou, and Papastavrou (2013) used Lego Technologies as a teaching tool for pre-service teachers and observed their development in both cognitive and non-cognitive domains. Pre-service teachers in the study demonstrated engagement with the Lego activities and were able "to overcome non-cognitive factors that often impede science learning outcomes such as gender issues, prior experiences, and attitudes toward science and self-esteem" (p. 627). Novak and Wisdom (2016) provided pre-service teachers with hands-on 3D printing design experiences to expose them to technology integration and to increase their engagement in science. Based on the analysis results they recommended 3D printing projects in elementary teacher science methods courses. Although studies exploring elementary teacher preparation in the use of technology are burgeoning, studies specific to science instruction are limited.

Few studies exist about elementary science teacher preparation for technology integration. However, scholars argue that "decisions about when to use technology, what technology to use, and for what purposes cannot be made in isolation of theories and research on learning, instruction, and assessment." (Lawless & Pellegrino, 2007, p. 581). This suggests that science teachers' preparation and their use of technologies are closely linked to goals, scientific inquiries, and assessments. Marino, Israel, Beecher, and Basham (2013) reported, after investigating middle school students' use of video games, that teachers who participated in the study thought that the video games linked to national science benchmarks, and assessments within games had more promise. Connecting technologies to science planning and instructional implementation and making more explicit technology integration efforts in teacher preparation are an area in need of more research. Kay (2006) concluded, for example, after reviewing 68 studies in technology integration in pre-service education, that collaboration among teacher preparation, including mentor teachers, is critical to improving abilities and attitudes towards the meaningful use of technology. More explicit, sustained, collaborative, and systematics efforts are needed to increase the chances for successful technology integration.

Discussion and Implications

The goal of this study was to understand the several phenomena of the professional journey of ETCs who were near the end of a teacher education program. Within the

context of a science methods course, this study examined ETCs' *learning science* and *learning-to-teach science* with a focus on a particular emerging technology, SL. The main curriculum, activities, and instructional materials of this evolving science methods course were designed to embody the concepts from Communities of Practice (Wenger, 1998) and Dialogic Inquiry (Wells, 2000) (See Figure 1).

The results of this study, which were compared and contrasted with the relevant literature reviews, revealed two gaps and three areas of growth, as shown in Figure 2. The first gap indicated the contrast between the positive learning experiences of the majority of the ETCs and the small number that stated they would use VR in their science teaching (Innovative integrators, 12%). The second gap pointed to a further contrast between the ETCs who noticed the positive impacts of using VR in their science teaching (both observant and innovative integrators, 71%) and their actual teaching practices, including lesson plans, peer teaching, and reflections, which showed little evidence of VR integration with science teaching.

Based on the gap analysis, we propose three areas of growth, which we identify as growth buds. Growth buds are where ETCs might develop their understanding of science and technology and grow to become science teachers. Growth bud 1 is related to providing ETCs with continuous opportunities to learn science and digital citizenship through emerging technologies. Growth bud 2 is associated with supporting ETCs to *see* themselves as innovative integrators in the context of teaching science and emerging technologies. Growth bud 3 involves strengthening ETCs' science teaching orientations (Friedrichsen et al., 2009) and teaching repertoires with emerging technologies.

In growth bud 1, the ETCs, as learners of science and emerging technology, considered the VR experience similar to inquiry-based learner-centered activities that follow the natural processes of scientific inquiry. Some ETCs believed that they had experienced a true form of self-discovery learning. Our data showed that the intense 3D visual objects and 3D science-related content motivated the ETCs to read more about scientific facts. Our data also indicated that the ETCs acquired digital skills while participating in the SL activity (e.g., learning how to use SL chats, e-notes, learning SL specific terms, learning which tool to use to teleport to islands). Therefore, the SL VR platform provided a unique learning experience for the ETCs who were projected as avatars in a 3D virtual world. Yet, some details found within these results require further consideration on how to guide ETCs to foster in-depth science learning and to cultivate mature digital citizenship along with various VR platforms or other emerging educational technologies (e.g., 3D printing technology, augmented reality, handheld digital devices).

In growth bud 2, the ETCs began to transform from their roles as undergraduate students to teachers who examine VR as an instructional material. As science teachers, 71% of the ETCs were willing to try integrating VR into their teaching. Yet, both the skeptical integrators (29%) and the observant integrators (59%) expressed great-to-moderate levels of negative attitudes towards using VR in their science teaching. This solidifies the findings from Russell et al.'s study (2003), where novice teachers were more reluctant to integrate technology into instruction although they had more technology knowledge and skills than experienced teachers.

Growth bud 2 is situated within this transitional space where ETCs shift from participating as learners of science to thinking like teachers. This is also the place where we expected that ETCs would channel their positive experiences as learners of VR into perceiving themselves as active users of VR to teach science. This process, however, was not as natural as expected for most participants. Few ETCs noticed affordances of VR for young learners and consequently did not see the merit of using VR in science instruction.

In growth bud 3, where ETCs engaged in designing and enacting science lessons, ETCs showed little interest in science phenomena represented in the 3D worlds and rarely engaged in discussing these phenomena in relation to goals, assessment or instructional strategies. This observation supports the findings from the existing literature, which indicate that teachers' perceptions towards technologies play a great role in their implementation of technologies. We conjecture that the negative attitudes that ETCs associated with VR affected their instructional decisions, which is evident in their 5E lesson plans and peer teaching. None of the ECTs used or integrated VR during their peer teaching. In addition, the use of other technologies was also limited. Only 29% of the 5E peer teaching teams incorporated technology, and only 4% of the ETCs mentioned any type of technology integration in their science notebooks. The technology they utilized was limited to computers, websites, or the Internet, which was used to research or find scientific facts for learners. In growth bud 3, ETCs need to further develop and strengthen pedagogical content knowledge. This is also the area for establishing, modeling, and practicing meaningful integration of technologies.

Overall, this study delineated ETCs' professional journey from learning science through VR to practicing innovation in science instruction. By analyzing the multiple types of data collected, interpreting the results, and reviewing the relevant literature, we were able to identify three phases of the journey and specific challenges that the ETCs faced at each phase of the journey (see Figure 2). Addressing the challenges identified in this study will help facilitate ETCs' professional journey into becoming successful elementary teachers.

The first challenge is that ETCs' positive learning experiences with VR did not naturally translate into willingness to use VR in science instruction. Skeptical integrators did not perceive themselves integrating VR into their science teaching. This could have resulted from the fact that the ETCs in the study were mostly first time VR users. Teacher preparation should address this lack of familiarity more explicitly and intently. As shown in Hughes's (2005) and others' development phases of technology integration, we learn that teachers go through the phases of familiarization and utilization before the integration phase. More sustained learning experiences in methods courses and possibly extended and collaborative

experiences in field placement will increase ETCs' knowledge and familiarity with emerging technologies.

Another challenge is situated in growth bud 3, where even the most advanced participants did not integrate VR into their science teaching. There are few studies on emerging technology integration within the context of elementary science methods courses; therefore, there was little that we could glean from existing literature on how to assist ETCs crossing this second challenge. Yet, based on our analysis of the results and the existing literature on general teacher preparation, we propose that we need to approach this challenge by systemizing ETCs' learning experiences and explicitly developing ETC's perceptions as teachers.

According to the International Society for Technology in Education, educators are positioned as learners, leaders, citizens, collaborators, designers, facilitators, and analysts. Also, learning science involves ETCs developing an in-depth understanding of the nature of science by actively exposing their misconceptions about the scientific worldview, scientific inquiry, and scientific enterprise (American Association for the Advancement of Science, 1990). During their 90-minute immersion in the VR world, the ETCs experienced a type of learning that was practical and exploratory. We consider this *first-hand positive experience* to be the very first requirement for the ETCs to be able to effectively fulfill their roles as learners, leaders, and citizens of the ICT.

However, the ETCs in our study did not maximize the experiences that this particular learning environment afforded. The majority of the ETCs had individualized experiences in the 3D environment. For example, they read about scientific facts from island to island, or interacted with the 3D contents either alone or with a small number of other avatars. Some ETCs were unsure of what to explore and how to explore; thus, they, as avatars, simply followed the other active avatars. The 3D world opens up hybrid space where ETCs might practice their identities as teachers of science. Yet, the significance of using avatars as projecting idealized selves was barely present in this study. In addition, they established only weak connections with their avatars for the purpose of practicing, visualizing, and internalizing professional identities. This link should be addressed in a more explicit manner in the methods course to help ETCs transition from a learner of science to a successful science teacher.

We also propose explicit and purposeful learner-centered pre- and post-VR activities. Initial VR activities may aim to engage ETCs to learn about technology, to learn about the features of the targeted technology, to learn how to present themselves as avatars, and the like. ETCs also need time to engage in activities that can develop connections with their projected-avatars (e.g., creating one's own avatar, collecting personal inventory items to express desired appearances and behaviors). After these pre-stages are explored, ETCs may be guided to explore how to maximize VR learning environments and how to identify science learning goals. During this phase, ETCs act-like, talk-like, and feel-like scientists and, therefore, become active learners of science. Post-activities should teach the ETCs how to meaningfully transfer their VR learning experiences into their teaching practices.

This study indicated that ETCs were not familiar with emerging technologies and faced challenges when designing and implementing science lessons that integrated emerging technologies. Yet, because of the unexpectedness and openness of the emerging technologies, VR provides a good place to learn about what it means to live in a digital world, to mature into a digital citizen, and to develop an appreciation for working within *ambiguous contexts*. Emerging technologies hold a high promise to enhance science learning and the ETCs' digital citizenship. Learning science is much more than just learning facts about science. A 3D world, such as SL, provides a unique learning space and a mode for future elementary science teachers to experience and practice science. Further, ETCs can safely design, practice, and experience science teaching in learner-centered, exploratory, and interdisciplinary environments and complete their journeys from learners to teachers.

Acknowledgements

We thank the elementary teacher candidates who participated in our study. We also thank Rhonda Berry, Sara Nelson and Jennifer Smith for their assistance during the process of making the collected data anonymous, systematized, and organized. We have benefited from Gauri Ramasubramanian for her yearlong assistance in cataloging the scanned data. Finally, we are grateful to Dr. Frankie S. Laanan, interim director of the Center for Excellence in Science, Mathematics, and Engineering Education at Iowa State University, and to Dr. John Hauptman for their support and encouragement.

References

Adler, P., & Adler, P. (1987). *Membership roles in field research.* Newbury Park, CA: Sage.

American Association for the Advancement of Science (AAAS). (1990). *Project 2061: Science for all Americans*. New York: Oxford University Press.

Applefield, J., Huber, R, & Moallem, M. (2001). Constructivism in theory and practice: Towards better understanding. The High School Journal, (Dec.-Jan.) 35-53.

Annetta, L., Mangrum, J., Holmes, S., Collazo, K., & Cheng, M. (2009). Bridging reality to virtual reality: Investigating gender effect and student engagement on learning through video game play in an elementary school classroom. *International Journal of Science Education*, *31*(8), 1091-1113.

Bang, E. (2013a). Exploring impact of the EED 420 Science Methods Course on pre-service elementary teachers views regarding the nature of science. *International Electronic Journal of Elementary Education*, *5*(3), 219-232.

Bang, E. (2013b). Hybrid-mentoring programs for beginning elementary science teachers. *International Journal of Education in Mathematics, Science and Technology*, *1*(1), 1-15.

Bakhtin, M. M. (1986). *Speech genres and other late essays*. (Y. McGee, Trans.). Austin: University of Texas Press.

Biological Sciences Curriculum Study (2011). Science teachers learning through lesson analysis (STeLLA). Retrieved from https://bscs.org/stella.

Bleicher, R. E. (2006). Nurturing confidence in pre-service elementary science teachers. *Journal of Science Teacher Education, 17*, 165–187.

Bourdieu, P. (1986) The forms of capital. In J. Richardson (Ed.) *Handbook of Theory and Research for the Sociology of Education* (New York, Greenwood), 241-258.

Burguillo, J. C. (2010). Using game theory and competition-based learning to stimulate student motivation and performance. *Computers & Education*, *55*(2), 566-575.

Bybee, R.W., Taylor, J.A., Gardner A., Scotter, P.V., Powell, J.C., Westbrook, A. & Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness. Retrieved from, https://bscs.org/sites/default/files/_media/about/downloads/BSCS_5E_Full_Report.pdf

California Council on Science and Technology. (2007). *Critical path analysis of California's science and mathematics teacher preparation system*. Retrieved from, http://www.google.com/url?sa=t&rct=j&q=&esrc=s&so urce=web&cd=4&ved=0CDkQFjAD&url=http://www.cftl.org/documents/2007/TCPA.pdf&ei=UAw3Upa_K6PO2wXspYDACQ&usg=AFQjCNGFf0lAjpq-zcb5oVpsBdiUnWq69A&bvm=bv.52164340,d.b2I

Chen, C. & Tsai, Y. (2012). Interactive augmented reality system for enhancing library instruction in elementary schools. *Computers & Education*, *59*(2), 638-652.

Corbin, J. M., & Strauss, A. L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications, Inc.

Costley, C., Elliott, G., & Gibbs, P. (2010). *Doing work based research: Approaches to enquiry for insider-researchers*. London, English: SAGE Publications Ltd.

Crowther, D. T., & Bonnstetter, R. J. (1997), *Science experiences and attitudes of elementary education majors as they experience Biology 295: A multiple case study and substantive theory*. Proceedings of the annual International Conference of the National Association for Research in Science Teaching (ERIC Document Reproduction Service No. ED 405220).

Dewey, J. (1997). *How We Think*. Dover Publications: Mineola NY.

diSessa, A. (2000). *Changing minds: Computers, learning, and literacy*. Cambridge, MA: MIT Press.

Dunne, C. (2011). The place of the literature review in grounded theory research. *International Journal of Social Research Methodology*, *14*(2), 111-124.

Dwyer, S. C., & Buckle, J. L. (2009). The space between: On being an insider-outsider in qualitative research. *International Journal of Qualitative Methods*, *8*(1), 54-63.

Engeström, Y. (1987). Comment on Blackler et al activity theory and the social construction of knowledge: a story of four umpires. *Organization—The Interdisciplinary Journal of Organisation, Theory and Society Studies*, *7*(2), May 2000, 301–310.

Friedman, T. L. (2005). *The world is flat: A brief history of the twenty-first century*. New York: Farrar, Straus and Giroux.

Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration? *Educational Technology Research and Development*, *53*(4), 25-39.

Ertmer, P. A., Addison, P., Lane, M., Ross, E., & Woods, D. (1999). Examining teachers' beliefs about the role of technology in the elementary classroom. *Journal of Research on Computing in Education*, *32*(1), 54-71.

Freire, P. (2003). *Pedagogy of the oppressed*. New York: Herder and Herder.

Friedrichsen P. J., Abell S. K., Pareja E. M., Brown P. L., Lankford D. M., Volkmann M. J. (2009). Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *Journal of Research in Science Teaching*, *46*(4), 357-383.

Gasson, S. (2004). Rigor in grounded theory research: An interpretive perspective on generating theory from qualitative field studies. In M. E. Whitman & A. B. Woszczynski (Eds.), *The handbook of information systems research* (pp. 79–102). Hershey, PA: Idea Group.

Glaser, B. G. & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New Brunswick, New Jersey: Aldine Transaction.

Glaser, B.G. (1998). *Doing grounded theory: Issues and discussions.* Mill Valley, CA: Sociology Press.

Greenfield, P.M. (1999). Historical change and cognitive change: A two-decade follow-up study in Zinacantan, a Maya community in Chiapas, Mexico. *Mind, culture, and activity*, *6*(2), 92-108.

Guba, E.G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries, *Educational Communication and Technology,* (29), 75–91.

Hadjiachilleos, S., Avraamidou, L., & Papastavrou, S. (2013). The use of Lego technologies in elementary teacher preparation. *Journal of Science Education and Technology, 22*(5), 614-629. DOI 10.1007/s10956-012-9418-4

Hermans, R., Tondeur, J., van Braak, J., & Valcke, M. (2008). The impact of primary school teachers' educational beliefs on the classroom use of computers. *Computers & Education*, *51*(4), 1499-1509.

Hew, K. F. & Brush, T. (2007). Integrating technology into K-12 teaching and learning: Current knowledge gaps and recommendations for future research. *Educational Technology Research Development*, *55*(3), 223-252.

Holland, D., W. Lachicotte, D. Skinner, and C. Cain. (1998). *Identity and agency in cultural worlds*. Cambridge, MA.: Harvard University Press.

Hughes, J. E. (2005). The role of teacher knowledge and learning experiences in forming technology-integrated pedagogy. *Journal of Technology and Teacher Education*, *13*(2), 277–302.

Inman, C., Wright, V. H., & Hartman, J. A. (2011). Use of Second Life in K-12 and higher education: A review of research. *Turkish Online Journal of Distance Education*, *12*(3), 67-85.

Jenkins, H., Clinton, K., Purushotma, R., Robinson, A. J., & Weigel, M. (2006). Confronting the challenges of participatory culture: Media education for the 21st century. Chicago, IL: MacArthur

Johnson, B., & Christensen, L. (2012). *Educational research: Quantitative, qualitative, and mixed approaches*. (4th ed.). Thousand Oaks, California: SAGE Publications Ltd.

Kay, R. H. (2006). Evaluating strategies used to incorporate technology into preservice education: A review of the literature. *Journal of Research on Technology in Education*, *38*(4), 385-410.

Kerawalla, L., Luckin, R., Seljeflot, S., & Woolard, A. (2006). "Making it real": Exploring the potential of augmented reality for teaching primary school science. *Virtual Reality*, *10*, 163-174.

Kim, P., (2006). Effects of 3D virtual reality of plate tectonics on fifth grade students' achievement and attitude toward science. *Interactive Learning Environments*, *14*(1), 25-34.

Kleiner, B., Thomas, N., Lewis, L., & Greene, B. (2007). *Educational technology in teacher education programs for initial licensure: Statistical analysis report*. (Master's thesis), Available from U.S. Department of Education. (NCES 2008-040). Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0CDoQFjAB&url=http://nces.ed.gov/pubs2008/2008040.pdf&ei=AQ83UrjfLoS22AX8wYHYCw&usg=AFQjCNFiLzihpzFPaDXUUk0dXf6JRhtewg&bvm=bv.52164340,d.b2I

Kuenzi, J. J. (2008). STEM (science, technology, engineering and mathematics) education: Background, federal policy and legislative action. CRS report for Congress. Retrieved from http://www.fas.org/sgp/crs/misc/RL33434.pdf

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. London: Cambridge University Press.

Lawless, K. A., & Pellegrino, J. W. (2007). Professional development in integrating technology into teaching and learning: Knowns, unknowns, and ways to pursue better questions and answers. *Review of Educational Research*. *77*(4), 575-614.

Liu, T. Y., & Chu, Y. L. (2010). Using ubiquitous games in an English listening and speaking course: Impact on learning outcomes and motivation. *Computers & Education*, *52*(1), 1-12.

Marino, M., Israel, M., Beecher, C. C., & Basham, J. D. (2013). Students' and teachers' perceptions of using video games to enhance science instruction. *Journal of Science Education Technology*, *22*(5), 667-680. DOI 10.1007/s10956-012-9421-9

Miles M. B., Huberman A. M. (1994). Qualitative data analysis: An expanded sourcebook (2nd ed.). Thousand Oaks, CA: Sage.

Moersch, C. (1995). Level of technology implementation (LoTi): A framework for measuring classroom technology use. *Learning and Leading with Technology*, *23*(3), 40-42.

Morris, M. W., Leung, K., Ames, D., & Lickel, B. (1999). Views from inside and outside: Integrating emic and etic insights about culture and justice judgments. *Academy of Management Review*, 24, 781–796.

National Academy of Sciences. (1996). *National science education standards*. Washington, DC: National Academy Press. Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0CEAQFjAD&url=http://www.csun.edu/science/ref/curriculum/reforms/nses/nses-complete.pdf&ei=oxE3UpWEBafF2AXi_oD4Cg&usg=AFQjCNF3hhWrH110sLjyTIubE8oFkwqi2Q&bvm=bv.52164340,d.b2I

National Academy of Sciences. (2010). Rising above the Gathering Storm, Revisited: Rapidly Approaching Category 5. Washington, DC: National Academies Press.

National Commission on Excellence in Education (NCEE). 1983. A nation at risk: The importance of educational reform. Retrieved from https://www2.ed.gov/pubs/NatAtRisk/risk.html

National Research Council. (2012). *A framework for k-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: The National Academies Press.

National Science and Technology Council. (2011). *The federal science, technology, engineering, and mathematics (STEM) education portfolio*. Retrieved from http://www.whitehouse.gov/administration/eop/ostp/nstc/committees/costem

National Science Board. (2006). *America's pressing challenge: Building a stronger foundation*. Available from National Science Foundation. (NSB0602). Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCkQFjAA&url=http://www.nsf.gov/statistics/nsb0602/nsb0602.pdf&ei=JhA3UpmxPKGA2AX-3YCoBw&usg=AFQjCNH-sqPs1bFIWBUmxNDiSBIVK-axYQ&bvm=bv.52164340,d.b2I

Newman, J. W., Abell, S. K., Hubbard, P. D., McDonald, J., Otaala, J., & Martini, M. (2004). Dilemmas of teaching

inquiry in elementary science methods. Journal of Science Teacher Education, *15*(4), 257-279.

Novak, E. & Wisdom, S. (2016, in press). Improving preservice elementary teachers' science engagement through 3D printing technology. *Journal of Immersive Education.*

Piburn, M., Sawada, D., Turley, J., Falconer, K., Benford, R., Bloom, I., & Judson, E. (2000). Reformed teaching observation protocol (RTOP): Reference manual (ACEPT Technical Report No. IN00-3). Tempe, AZ: Arizona Collaborative for Excellence in the Preparation of Teachers. *Measuring Reform Practices in Science and Mathematics Classrooms: The Reformed Teaching Observation Protocol*. Retrieved from http://physicsed.buffalostate.edu/AZTEC/rtop/RTOP_full/PDF/RTOP_ref_man_IN003.pdf

Reed, J. and Procter, S. (eds) (1995) *Practitioner research in health care:The inside story,* London: Chapman and Hall.

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.

Russell, M., Bebell, D., O'Dwayer, L., & O'Connor, K. (2003). Examining teacher technology use: Implications for preservice and inservice teacher preparation. *Journal of Teacher Education*, *54*(4), 297-310.

Shenton, A.K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, *22*, 63-75.

Slotta, J. & M. Linn, M. (2009), *WISE science: Web-based inquiry in the classroom* NY: Teachers College Press.

Srivastava, P., & Hopwood, N. (2009). A practical iterative framework for qualitative data analysis. *International Journal of Qualitative Methods*, *8*(1), 76-84.

Sun, K., Lin, C., & Wang, S. (2010). A 3-D virtual reality model of the Sun and the Moon for e-learning at elementary schools. *International Journal of Science and Mathematics Education*, *8*(4), 689-710.

Sung, H. & Hwang, G. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, *63*, 43-51.

Toledo, C. (2005). A five-stage model of computer technology integration into teacher education curriculum. *Contemporary Issues in Technology and Teacher Education*, *5*(2), 177-191.

Triangle Coalition for STEM Education (2017, March 26). Math and science education funding. Retrieved from www.trianglecoalition.org/president-calls-for-more-stem-teachers-grads.

U.S. Congress Joint Economic Committee. (2012). *STEM education: Preparing for the jobs of the future*. Retrieved from http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CCsQFjAA&url=http://www.jec.senate.gov/public/index.cfm?a=Files.Serve&File_id=6aaa7e1f-9586-47be-82e7-326f47658320&ei=TRE3UpCjNOTF2QWMs4D4CA&usg=AFQjCNGonTJ74BDEYSZ4kofuZK5faaMnPw&bvm=bv.52164340,d.b2I

Veal, W. R., & MaKinster, J. G. (1999). *Pedagogical content knowledge taxonomies*. Electronic Journal of Science Education, *3*(4): Retrieved from http://wolfweb.unr.edu/homepage/crowther/ejse/vealmak.html

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* Cambridge, Mass.: Harvard University Press.

Wallerstein, I. (1998) Utopistics, or Historical Choices of the Twenty First Century. New York, NY: New Press.

Wells G. (2000). Dialogic inquiry in education. In Lee C., Smagorinsky P. (Eds.) Vygotskyian perspectives on literacy research (pp. 51–85). New York: Cambridge University Press.

Wenger, E. (1998) *Communities of practice: learning, meaning, and identity.* Cambridge University Press.

Wu, H-K., Lee, S. W-Y., Chang, H-Y., Liang, J-C. (2013). Current status, opportunities and challenges of augmented reality. *Computers & Education*, *62*, 41-49.

Appendix

A description of a virtual reality activity embedded in the science methods course

QUEST: *We Are SciTeachers!*

1. Establish your dream team—a group of three or four: The goal of this quest is to simply play with Second Life (SL), with your dream team, and see how your group can solve the quests below within a scheduled time period—90 minutes. In the meantime, I will be your facilitator. If you have any questions during the activity, feel free to IM to my avatar.

    - Full SL name: evesapple75 Kuhn (a whole class facilitator)

2. Assign a role to each team member: You will need a *Facilitator* who will lead your team, a *Recorder* who will record game strategies and tactics, and a *Photographer* who will take snapshots as your team explores selected islands. Short activity reports in Word Document, will also be required as documentation.

3. Before you meet in the virtual world, brainstorm how your team will finish this quest or how it will explore science-related islands in time—by developing some basic strategies, rules, confirming roles, and tasks for the quest. For instance, you can make rules like the following: Each group member must send one sentence that captures the experience on the SL Orientation Island to your SL recorder. When the team is on an island, each team member will move together with the others. *Or,* each team member may explore the island individually, and teleport team members—if he or she finds interesting sites within the island.

4. Start your VR activity by offering friendship to your group members and the instructor. Now, you can IM, and teleport each other to different islands and places.

5. Open up the PDF document entitled *Science-Related Places in Second Life* from our Blackboard, under "Virtual Reality Science Field Trips." See what islands are available, and evaluate where your dream team wants to explore.

    - By using the SEARCH button, teleport to at least **five science-related islands** on the list, *as a team*, and explore. Make sure that you open up a mini-map at each island and use it (the yellow dot is you; and the green dots are others).

    - Recommend Islands from the list: Euclidia Space Planetarium, The Abyss Observatory, Synthetic Biology Interactive, Elucian Islands, MICA, Energy Island, the Neil A. Armstrong Library & Archives, Genome Island, Fire Sabre's Volcano, the Jet Propulsion Laboratory, Virtual Hallucinations, Science Friday, The Bell of the Firmament, Cloud Chateau, David Rumsey Maps, Astronomy 2009 Island, Apollo Moon Landing, Virtual Stomach Museum, Tech Museum of Innovation, and the Spaceflight Museum.

6. Upload your Word Document to the VR table on our wiki page. Make sure that you click the SAVE button at the bottom of the wiki page when you are finished—one document per one group. Also, send this same document to your team members so you have multiple backup files.

7. Individually, answer the following questions in Word Document and upload it to our class PBwiki.

    a. What skills did you learn in terms of science content, vocabulary, or inquiry skills?
    b. How did you grow in your confidence with, and abilities in, digital skills and virtual reality?
    c. What do you want to report about your experiences with virtual reality, particularly regarding its potential(s) for science education?
    d. How do you think you might incorporate virtual reality—and other similar virtual worlds—for science education activities in the future in your classrooms?

Congratulations!

This page is intentionally left blank

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effects of Using Concept Cartoons in Astronomy Subjects on Critical Thinking Skills among Seventh Grade Student*

Filiz Demirci [a,*], Cengiz Özyürek [b]

Received: 20 September 2017
Revised: 3 November 2017
Accepted: 12 November 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236119

Abstract

The objective of this study is to research the effects of using concept cartoons in the "Solar System and Beyond" unit, which is included in seventh grade science lessons, on students' critical thinking skills. The study group consisted of 58 students, selected using an appropriate sampling method, who were students in a state secondary school, which is close to the city centre. The study used a pre-test and post-test matched control group design. During the three-week-long experimental teaching process, the students in the control group (n=30) were taught according to the curriculum of science lessons, while concept cartoons developed by the researchers were used with the students in the experimental group (n= 28). The Cornell Critical Thinking Test-Level X (CCT-X), as developed by Ennis & Millman (1989), was used as the data collection tool. The results of the study showed decreases in the critical thinking skills scores of the students in the control group, while a statistically significant difference was found in the critical thinking skills of the students in the experimental group. In addition, a statistically significant difference was observed in favour of the experimental group when comparing the CCT-X scores of the experimental group's use of concept cartoons and those of the control group's used existing teaching methods.

Keywords: Solar system and beyond, astronomy education, critical thinking, concept cartoons.

Introduction

Astronomy is a field of science that gives us historical information about humankind and data on the structure and history of the universe in order to understand what kind of a future the universe will have, as well as the phenomena taking place in the sky. In short, it is a field of science that is interested in every foundation stone of the universe, from the smallest to the biggest. As one of the oldest sciences, astronomy offers individuals scientific thinking skills and plays a role in their capacity to make judgements about the universe (Ministry of Education [MoNE], 2010). Astronomy enables human beings to understand the universe. It is necessary for human beings to make deductions about events and phenomena by using the information they have, their logic, scientific methods and critical thinking skills in order to be able to gather scientifically acceptable information about the world and they universe they live in. It is especially important for individuals to use critical thinking skills in order to be able to find out which of the existing assertions available to believe in.

Ennis (1985, p. 45) defined critical thinking as "reasonable and reflective thinking which focuses on what one believes in or what to do". Evancho (2000) defined critical thinking as the ability of "the individual to make analytic and assessment-oriented conscious judgments and express these judgments to reach a decision as to that s/he shall do or believe" (quoted in Seferoğlu & Akbıyık, 2006: 195).

According to Ennis (1991), critical thinking includes the following features:

1. *Clarity:* disposition to being clear about the meaning of what is expressed or written
2. *Focus:* disposition to finding out and continuing to focus on the question or the result
3. *Total situation:* disposition to consider the situation in all its aspects
4. *Reasons:* disposition to look for and present reasons to justify logical reasoning
5. *Trying to be well informed:* disposition to present proof about a subject
6. *Alternatives:* disposition to look for different ways to understand a situation
7. *Precision:* disposition to try to obtain definite information to the extent permitted by the subject
8. *Self-awareness:* disposition to try to be self-aware about each of one's primary needs
9. *Open-mindedness*: disposition to be open-minded by thinking of different perspectives other than one's own
10. *Caution:* disposition to refrain from making an important judgement when evidence and reason are not sufficient
11. *Non-scepticism:* disposition to have a position or change it when evidence and reasons are sufficient
12. *Using one's abilities:* disposition to foster and use critical thinking skills

According to educational philosophy, critical thinking is an indispensable part of education, rather than simply being

[a,*] Corresponding author: Filiz Demirci, Department of Mathematics and Science Education, Ordu University, Ordu, Turkey. E-mail: filizdemirci@odu.edu.tr
[b] Department of Mathematics and Science Education, Ordu University, Ordu, Turkey. E-mail: cengizozyurek@odu.edu.tr

another option (Norris, 1985). Today, around the world, it is widely accepted that a society that is not dominated by critical thinking cannot be a developed society (Aybek, 2007). As the future leaders or decision makers in our society, today's students should be able to think openly, with a sceptical and questioning attitude when deciding on social problems, while being able to assess the arguments of an issue critically and reach healthy decisions (Tümay & Köseoğlu, 2011).

Critical thinking enables students to synthesize, analyse and assess students' thoughts, beliefs and actions in order to guide them better. Critical thinking has become both a result and a necessity for human beings who are experiencing fast changes and who need and desire to predict the future and prepare for it ideationally (Aybek, 2007).

Besides, critical thinking and critical thinking skills can be earned and can be developed through education (Schafersman, 1991). Today, critical thinking, which is at the forefront of learning and teaching thinking skills, is an important factor in terms of increasing students' academic achievement, encouraging them to produce information, enabling them to learn by being less dependent on textbooks and teachers, helping them to assess social structures and encouraging them to change these structures (Anonymous, 2016). Thus, learning and teaching critical thinking should have an important place. Tsai, Chen, Chang and Chang (2013) concluded that developing students' critical thinking skills helped them to comprehend information delivered in science lessons. A great number of previous studies on this subject shows that students' critical thinking skills and their academic performance/achievement are closely related. In other words, students with higher critical thinking skills are more successful in lessons (Ip, Lee, Lee, Chau, Wootton & Chang, 2000; Giancarlo & Facione, 2001).

Browne & Freeman (2000) stated that critical thinking can be taught to students with their active participation. One of the teaching techniques that encourages students participate actively in class is the use of concept cartoons, which gives them a discussion environment. Concept cartoons can be used as a teaching tool to encourage students to think critically and participate in discussions (Cho & Reich, 2008).

Along with learning scientific knowledge, enjoying the lesson is also a goal of science education. Within this context, humorous teaching materials can make science lessons more fun and thus help students to enjoy science (Özdemir, 2017). In addition, according to Percy (2006), studies have shown that students who are interested in the teaching materials learn more effectively, while Jiwaji (2016) stated that astronomy encouraged students to increase their interest in science. In light of all these statements, while taking into consideration the principle that "students learn more when they are interested in the subject or when they are motivated", it is important to use concept cartoons developed for astronomy subjects.

Concept cartoons represent a method that can be used as a learning, teaching and assessment technique in science teaching (Bahrani & Soltani, 2011; Chin & Teou, 2009; Dabell, 2008; Kabapınar, 2005; Keogh, Naylor & Wilson, 1998; Keogh & Naylor, 1999; Naylor & Keogh, 2010, 2012). Dabell (2004) expressed the advantages of concept cartoons when used as teaching tool as follows:

- They give students the opportunity to question their ideas, to extend their thoughts and to bring different perspectives to events.
- They help to create a discussion environment and active participation of students.
- They guide students in solving the problems they encounter in their daily lives.
- They reveal the uncertainties and misconceptions in students' minds.
- They can be used in activities conducted to combine or extend subjects (quoted in Balım, İnel & Evrekli, 2008).

Regarding Dabell's expressions, it can be said that a well-formed concept cartoon will contribute to students' mental development, especially to their critical thinking skills. In addition, Browne & Freeman (2000), classes in which critical thought is dominant generally have four characteristics: "frequently asked questions", "growing tension", "exposure with the possibility of reached conclusions" and "encouraging active learning". Bahrani and Soltani (2011) stated that, by using concept cartoons: 1. students' attention can be engaged by the words and pictures in the concept cartoon, 2. the motivation and attitude to learn can increase positively, 3. they are useful and alternative tools for the cognitive development of students, 4. they help to develop students' higher order critical thinking skills, 5. they represent a starting point to make students think about their ideas or feelings about a situation, and 6. they help to create a discussion environment in the classroom.

This instructional technique can benefit students by means of motivating effects of humor and creative drawing, along with mental manipulations (Zousel, Rule & Logan, 2013). Presenting a discussion environment with student-centred teaching plays an effective role in developing students' critical thinking skills (Alkaya, 2006). Besides enabling students to think freely and express themselves or their ideas in writing or orally, concept cartoons also develop students' critical thinking skills (Özlap, 2006). Within this context, Özalp (2006) stated that, during the process of interpreting the cartoons, students make connections with the drawings, create stories in their minds by interpreting these connections and think critically by using this method.

A great number of studies in which the use of concept cartoons is approached as a teaching method has shown that concept cartoons encourage students to think critically and create critical discussions by encouraging them to participate in the lesson (Balım et al., 2008; Kabapınar, 2005; Keogh & Naylor, 1999; Keogh, Naylor, de Boo & Feasey, 1999). Kabapınar (2005) concluded that using concept cartoons as a teaching method increased students' participation in lessons, facilitated critical discussion and encouraged them to research the validity of the thoughts presented. Balım et al. (2008) studied the effect of concept cartoons in science education on students' academic achievement and questioning-based

learning skills, concluding that using concept cartoons was effective in developing students' critical thinking skills.

When the literature on the teaching of astronomy subjects is reviewed, it can be seen that, rather than being confined to the teaching of conceptual definitions, it is necessary to use observations, discussions and modelling, for example, to help the students make interconceptual relations in a meaningful way (Diakidoy & Kendeou, 2001; Zeilik et al., 1997). Percy (1998) stated that, although astronomy is taught extensively in primary education, teachers have little or no information about astronomy, while the teaching techniques used are generally confined to coursebooks. In addition, according to the author (1998), in many studies conducted on astronomy teaching, it was claimed that telescopes and computers are not necessary, while hands-on activities, which are cheaper, are recommended as being more effective. On the other hand, astronomy education plays an important role in developing the critical thinking skills of individuals (Uçar & Demircioğlu, 2011). In the light of all these views, the objective of this study is to show whether a teaching method in which concept cartoons are used for teaching astronomy subjects, which include several abstract concepts, will result in a significant difference in students' critical thinking skills. In addition, when the literature was reviewed, no studies were found that discussed the effect of using the concept cartoon technique in the "Solar System and Beyond" unit of seventh grade science lessons on students' critical thinking. It is thought that this study will determine the effectiveness of the concept cartoon technique on students' critical thinking skills and present ideas to teachers, researchers and programme makers.

Problem Statement

This study basically tries to answer the question: "What are the effects of using the concept cartoon technique in the "Solar System and Beyond" unit in the "Earth and Universe" subject of seventh grade science lessons on students' critical thinking skills?"

Subproblems

In the "Solar System and Beyond" unit of seventh grade science lessons;

- Are there statistically significant differences in the The Cornell Critical Thinking Test, Level X (CCT-X) pre-test and post-test scores for the control group students taught with the existing teaching method?
- Are there statistically significant differences in the CCT-X pre-test and post-test scores for the experimental group students used the concept cartoon?

Are there statistically significant differences between the CCT-X pre-test scores for the control group students taught with the existing teaching method and the experimental group students used concept cartoons?

Methodology

Study Design

The quasi-experimental design was used in this study. Quasi-experimental design is a type of design in which subjects are randomly assigned to different experimental and control groups. The most important advantage of this design is the possibility of randomly assigning existing groups into different groups when it is impossible to randomly assign, especially in the social sciences (Fraenkel et al., 2012). The design of this study is given in Table 1.

Table 1. *Pre-test and Post-test Semi-Experimental Design with Control Group*

Group	Pre-test	Teaching Method	Post-test
Experimental	CCT-X	Concept Cartoon Technique	CCT-X
Control	CCT-X	Existing Teaching Methods	CCT-X

Study Group

The study was conducted in a state secondary school close to the city centre with a middle socio-economic level. The sample was determined using an appropriate sampling method among the state schools in the city of Ordu, which was regarded as the universe of this study. Two classes were chosen randomly from seventh graders in the school. The experimental group consisted of 28 students (17 girls, 11 boys), while the control group consisted of 30 students (13 girls, 17 boys). Seventh graders were chosen as the study group since the unit "Solar System and Beyond", in which astronomy subjects are taught to a great extent, is taught at this grade.

Data Collection Tool

The CCT-X, developed by Ennis & Millman (1989) for between fourth and 14th grade was used as the data collection tool. The test has a total of 71 multiple choice questions and five subdimensions (induction, deduction, making assumptions, making observations and questioning the credibility of sources). The test was adapted into Turkish by Akar, Yüve and Acun, while the Cronbach alpha reliability coefficient value of the test was detected as 0.71 (quoted in Akar, 2007). The data were collected using the CCT-X as a pre-test for the control and experimental groups before the experimental process and as a post-test for both groups after the experimental process. As a result of the reliability analysis of this sample group, the Cronbach alpha reliability coefficient of the test was found to be 0.79.

Developing Concept Cartoons

Concept cartoons, which were used as teaching tool in the experimental group, were prepared by the researchers, based on the content of subjects/concepts in the "Solar System and Beyond" unit of seventh grade science lessons (MoNE, 2013). Common cartoon characters and the Scratch[1] program were used to create the concept

[1] *A free program in which users can create interactive students, games and animations of their own.*

cartoons. A total of 22 concept cartoons (three on space and the universe, one on light years, one on star clusters, four on stars, four on the solar system and planets, two on comets, one on meteors, one on galaxies, one on telescopes, one on space pollution, one on astronomers/astronauts, and two on astronomy/astrology concepts) were created, with the aim of developing students' critical thinking skills.

The prepared concept cartoons were examined by a lecturer in the Science Education Department and a professor in the Astronomy and Space Sciences Department in terms of their conceptual framework and the scientific expressions in the speech balloons. The experts stated that all of the concept cartoons (without any exclusion) could be used in the pilot study after the recommended editing was made. In addition, care was taken to create the concept cartoons in a way that they would help to develop students' critical thinking skills. For this purpose, the researchers tried to create at least three speech balloons (characters), which included different ways of thinking with content, in order to facilitate a discussion environment in the classroom. An example of the concept cartoons used in experimental teaching process is given in Appendix.

Pilot Study

Before the study was conducted, the required permissions were sought from the Ordu Provincial Directorate for National Education, from the competent ethical board and from the directorate of the school in which the study was to be conducted in order to realize the legal procedures. In addition, oral consent was obtained from each teacher and student to confirm their willingness to participate in the study voluntarily.

The study was completed in four weeks and ran as follows: informing the students and teachers about the concept cartoon technique, the pilot study, pre-tests, experimental processes and post-tests. All the applications in the study were conducted in the classroom environment. Firstly, the students in the experimental group and the teachers were informed about the application process involving concept cartoons before the teaching was carried out, after which the pilot study was conducted with the cartoons in order to determine the comprehensibility of the cartoons by the students, as well as the estimated application period of a cartoon. Within this context, two weeks before the experimental application, the lesson plan and 22 concept cartoons were provided to the teacher. In addition, informative interviews were made with the teacher about the experimental application outside the learning-teaching process. The pilot study, which lasted for a week (four hours of class), was delivered to the students in the experimental group on the subject of the "transformation of electrical energy" by using the cartoon concept. After the researchers observed each lesson in the pilot study without participating, they gave feedback to the teacher or students about the problems and shortcomings that they noticed about the application.

Experiment

One week before the experiment was conducted, the CCT-X was given to both groups as a pre-test by the science teacher. The teaching, which lasted for three weeks (12 hours of class) in both groups, was conducted in line with the existing curriculum for the control group, while concept cartoons were used for the experimental group. After the experiment was conducted, the CCT-X was given to both groups as a post-test by the science teacher. During this process, the researcher made no interventions in the classroom.

When the teacher came to class, he started by projecting the concept cartoon on the board so that all the students could see easily it. The teacher introduced the characters and the thoughts they advocated to students. After this, the teacher asked which characters' thoughts the students agreed with and why. All the students in the classroom were given the chance to express their ideas, which were or were not in the concept cartoon, openly and easily. Thus, the students were given the chance to be aware of different views, as well as question their own knowledge and views. Following a class discussion, the teacher invited the students to suggest a method to research the correctness of the thoughts they advocated. During this process, while the students were researching the correctness of their thoughts, the teacher offered no scientific explanations about the concept. At the end of the method, the students discussed their research results in the classroom with the teacher and reached a conclusion.

Data Analysis

The data obtained with the CCT-X in both groups were sent to the company that had the copyright for using the test in Turkey, which calculated the total scores for each student taking the test. The SPSS 22.0 program was used in the analysis of the data.

The t-test was used to find out whether there were significant differences between the scores for the control and experimental groups. Firstly, the paired sample t-test was used to find out whether there were significant differences in the pre-test and post-tests scores for the groups, while an independent samples t-test was used to find out whether there were significant differences between the groups (Field, 2009). For the analysis of the data, Shapiro-Wilk normality test values (since the group size was less than 50) were used to test the normality assumption, while the Levene test results were examined to test the equality assumption of group variances. In the analysis of the data for the first and second subproblems, which showed a normal distribution and equal group variance, the dependent groups t-test was used. In the two-way ANOVA for mixed measurements, there are two different factors to successively measure in both groups (experimental and control) using the pre-test and the post-test. By carrying out this analysis, it can be determined whether: 1. there is a significant difference between the total pre-test and post-test scores of both groups measured, irrespective of the difference between successive measurements, 2. there is a difference between successive measurements of all participants,

irrespective of groups, and 3. the change between the pre-test and post-test scores is significantly different in terms of the experimental and control groups (Can, 2014: 247).

In addition, the effect size (partial eta squared value) of the data, showing a significant difference in the analyses, was calculated. The effect size indicates how much the independent variable explains the total variance in the dependent variable (Field, 2009). It can be said that the partial eta squared (partial η^2). The level of significance (p) was taken as 0.05.

Results

Results of the First Subproblem

The data obtained from the subproblem, *"Are there statistically significant differences in the CCT-X pre-test and post-test scores for the control group students taught with the existing teaching method in the 'Solar System and Beyond' unit of seventh grade science lessons?"*, were analysed with the dependent groups t-test. The descriptive results of the students in the control group for the CCT-X are given in Table 2.

Table 2. *Descriptive Analysis Results of the CCT-X Pre-test and Post-test Scores for the Control Group*

Application	*N*	Min.	Max.	*M*	sd	Variance	Skewness	Kurtosis
Pre-test	30	21.00	52.00	38.27	6.67	44.55	-0.50	0.83
Post-test	30	19.00	52.00	37.57	7.01	49.15	-0.34	0.83

The lowest score one can get from the CCT-X is 0, while the highest score is 71. The lowest pre-test score for the students in the control group was 21.00, while their highest pre-test score was 52.00; and their lowest post-test score was 19.00, while their highest post-test score was 52.00. When Table 2 is examined, it can be seen that the post-test mean scores for the students in the control group ($M_{post\text{-}test}$= 37.57) are relatively lower when compared with their pre-test score means ($M_{pre\text{-}test}$= 38.27). In addition, it can be seen that the skewness and kurtosis values of the normality distribution of the pre-test and post-test scores for the control group students were between -1 and +1 (Table 2). However, normality test results should also be examined to determine whether the data of the control group are normally distributed. Normality test results of the CCT-X pre-test and post-test for the control group are given in Table 3.

When Shapiro-Wilk values are examined, it can be seen that the CCT-X pre-test and post-test results for the control group are normally distributed (p>0.05) (Table 3). In addition, when the skewness and kurtosis values in Table 2 were assessed with the values determined by the normality test, it was found that the parametric tests were suitable for the analysis of the pre-test and post-test score means of the control group.

Table 3. *Normality Test Results of the CCT-X Pre-test and Post-test for the Control Group*

Application	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	sd	*p*	Statistic	sd	*p*
Pre-test	0.14	30	0.12	0.97	30	0.52
Post-test	0.13	30	0.17	0.97	30	0.51

Thus, the independent groups t-test was used to find out whether there were statistically significant differences between the CCT-X pre-test and post-test score means of the control group. The independent groups t-test analysis results of the pre-test and post-test scores for the control group are given in Table 4.

Table 4. *t-test Results of the CCT-X Pre-test and Post-test for the Control Group*

Test	Application	*N*	*M*	sd	df	*t*	*p*
CCT-X	Pre-test	30	38.27	6.67	29	0.87	0.40
	Post-test	30	37.57	7.01			

When Table 4 is examined, it can be seen that there is no statistically significant difference between the pre-test and post-test scores for the students in the control group. It can be seen that the CCT-X post-test score means ($M_{post\text{-}test}$= 37.57) are lower when compared with the pre-test score means ($M_{pre\text{-}test}$= 38.27). This result can mean that the teaching process conducted with the existing teaching methods has no positive influence on increasing the critical thinking test scores for the students.

Results of the Second Subproblem

The data obtained from the subproblem, *"Are there statistically significant differences in the CCT-X pre-test and post-test scores for the experimental group students used the concept cartoon in the 'Solar System and Beyond' unit of seventh grade science lessons?"*, were analysed with the dependent groups t-test. The descriptive results of the students in the experimental group for the CCT-X are given in Table 5.

Table 5. *Descriptive Analysis Results of the CCT-X Pre-test and Post-test Scores for the Experimental Group*

Application	*N*	Min.	Max.	*M*	sd	Variance	Skewness	Kurtosis
Pre-test	28	19.00	42.00	31.10	5.38	28.91	-0.30	0.03
Post-test	28	21.00	44.00	34.50	6.88	47.37	-0.54	-0.91

The lowest score one can get from CCT-X is 0, while the highest score is 71. It can be seen that the lowest pre-test score for the students in the experimental group was 19.00, while their highest pre-test score was 42.00; and their lowest post-test score was 21.00, while their highest post-test score was 44.00. When Table 5 is examined, it can be seen that the post-test mean scores for the students in the experimental group ($M_{post\text{-}test}$= 34.50) are relatively higher when compared with their pre-test score means ($M_{pre\text{-}test}$= 31.10). In addition, it can be seen that the skewness and kurtosis values of the normality distribution of the pre-test and post-test scores for the experimental group students were between -1 and +1 (Table 5). However, the normality test results should also be examined to interpret whether the data of the experimental group are normally distributed. The normality test results of the CCT-X pre-test and post-test for the experimental group are given in Table 6.

Table 6. *Normality Test Results of the CCT-X Pre-test and Post-test for the Experimental Group*

Application	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	sd	*p*	Statistic	sd	*p*
Pre-test	0.11	28	0.20	0.98	28	0.80
Post-test	0.18	28	0.02	0.92	28	0.06

When Shapiro-Wilk values are examined, it can be seen that the CCT-X pre-test and post-test results of the experimental group are normally distributed (p>0.05) (Table 6). In addition, when the skewness and kurtosis values in Table 5 were assessed with the values determined by the normality test, it was found that parametric tests were suitable for the analysis of pre-test and post-test score means for the experimental group. Thus, the independent groups t-test was used to find out whether there were statistically significant differences between the CCT-X pre-test and post-test score means for the experimental group. The independent groups t-test analysis results of the pre-test and post-test scores for the experimental group are given in Table 7.

Table 7. *T-test Results of the CCT-X Pre-test and Post-test for the Experimental Group*

Test	Application	*N*	*M*	sd	df	*t*	*p*
CCT-X	Pre-test	28	31.11	5.38	27	-2.07	0.04*
	Post-test	28	34.50	6.88			

*: p<0.05

When Table 7 is examined, it can be seen that there is a statistically significant difference between the pre-test and post-test scores for the students in the experimental group in favour of the post-test scores ($t_{(27)}$= -2.07, p≤0.05). It can be seen that the CCT-X post-test score means ($M_{post\text{-}test}$= 34.50) are relatively higher when compared with the pre-test score means ($M_{pre\text{-}test}$= 31.11). This result could indicate that the teaching process conducted with the concept cartoons significantly increased the critical thinking test scores among students.

Results of the Third Subproblem

The data obtained for the third subproblem, *"Is there statistically significant difference between the CCT-X pre-test and post-test scores for the control and experimental group students used concept cartoons in the 'Solar System and Beyond' unit of seventh grade science lessons?"*, was analysed with the two-way ANOVA for mixed measures. Before the two-way ANOVA for mixed measures, primary hypotheses were tested. According to the first hypothesis, the dependent variable was found to be within the least interval scale. Since the CCT-X is a scale with equal intervals, this hypothesis was met. For the second hypothesis, the skewness and kurtosis values in Table 8 and the Shapiro-Wilk normality test values in Table 9 were analysed. According to the results obtained, it can be said that the skewness and kurtosis values were between +1 and -1, while the Shapiro-Wilk values met the normal distribution assumption; that is, the CCT-X scores met the normal distribution assumption in each subgroup. For the third hypothesis, the homogeneity of the variance for scores from the CCT-X of both groups was tested with Levene's test. It was found that the Levene's test result met the homogeneity of the variance in students' pre-test scores [$F_{(1,56)}$= 0.81, p>0.05] and post-test scores. Thus, assumptions were met for the two-way ANOVA analysis. As a result of the Box's test conducted for the fourth assumption, no significant difference was found between the paired combinations of measurement variances [$F_{(3,672164.56)}$= 2.26, p>0.05]. As a result of Mauchly's Test of Sphericity conducted for the fifth assumption, it was found that the difference score in repeated measurements of the participants in any group differed significantly from difference scores for the other participants. Thus, all the assumptions of the two-way ANOVA for mixed measures were met. The descriptive analysis results of the CCT-X pre-test and post-test scores for students in the experimental and control groups are given in Table 8.

Table 8. *Descriptive Statistic Analysis Results Related to the CCT-X Pre-test and Post-test Scores*

Application	Group	*N*	Min.	Max.	*M*	sd	Variance	Skewness	Kurtosis
Pre-test	Control	30	21.00	52.00	38.11	6.88	47.28	-0.43	0.39
	Experiment	28	19.00	42.00	31.11	5.38	28.91	-0.30	0.03
Post-test	Control	30	19.00	52.00	37.39	7.19	51.66	-0.28	0.58
	Experiment	28	21.00	44.00	34.50	6.88	47.37	-0.54	-0.91

Table 9. *Normality Test Results of the CCT-X Pre-test and Post-test Mean Scores for the Experimental and Control Groups*

Application	Group	Shapiro-Wilk		
		Statistics	df	*p*
Pre-test	Control	0.98	28	0.74
	Experiment	0.98	28	0.79
Post-test	Control	0.97	28	0.51
	Experiment	0.92	28	0.06

In the study, it can be seen that the pre-test mean score for students in the experimental group, in which concept cartoons were used, was 31.11. Although their post-test mean score increased to 34.50 after the application, it can be seen that the pre-test mean score for students who were taught with the present application was 38.27, while their post-test mean scores decreased to 37.57 after the application (Table 8 and Figure 1). According to the results, it was found that there was an increase in the CCT-X scores for the experimental and control group students in favour of the experimental group following the experiment; however, the decrease in the CCT-X post-test scores for control group students was remarkable (Figure 2).

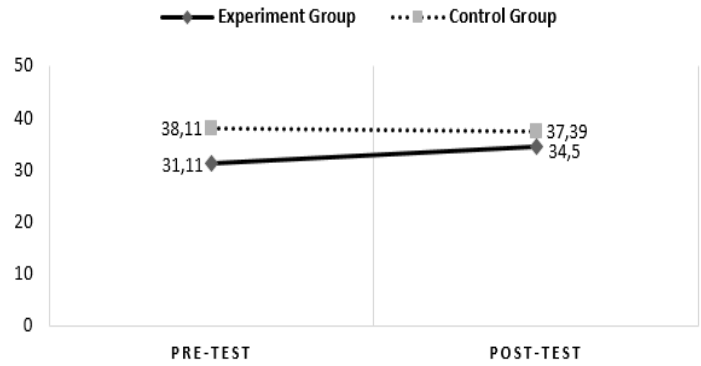


Figure 1. *Comparison of the Results of the CCT-X Pre-test and Post-test Mean Scores for the Experimental and Control Groups*

The results of the two-way ANOVA for mixed measures about whether there was a significant difference between the CCT-X pre-test and post-test mean scores for the experimental and control group students are given in Table 10.

Table 10. *Results of the Two-way ANOVA for Mixed Measures of the CCT-X Pre-test and Post-test Scores for the Experimental and Control Groups*

Source	Sum of Squares	df	Mean Square	*F*	*p*	η^2	Observed Power
Between Groups		57					
Group (CC/PA)	757.267	1	757.267	12.20	0.00*	0.18	0.93
Error	3475.423	56	62.061				
Within Groups		58					
Measurement (Pre-test/Post-test)	52.511	1	52.511	2.26	0.14	0.04	0.32
Group*Measurement	**121.304**	**1**	**121.304**	**5.22**	**0.03***	**0.09**	**0.61**
Error	1301.489	56	23.241				
Total		115					

*$p<0.05$; CC: Concept Cartoons; PA: Present Application (Using the 2013 Science Curriculum)

An analysis of the two-way ANOVA for mixed measures was conducted in order to find out the effect of teaching with concept cartoons on the critical thinking skills of seventh graders. When Table 10 is examined, it can be seen that the critical thinking skills of the experimental group, in which concept cartoons were used, and the control group, which was taught with the present application, differed significantly between the pre-test and post-test stages, that is, the combined effect of being in the experimental and control groups and repeated measurement factors was found to be statistically significant ($F_{(1,56)}$= 5.22, p<0.05).

In this study, primary effect tests were also used to test the effect of teaching methods applied in the experimental and control groups on students' critical thinking skills. When Table 9 is examined in terms of the group primary effect, a statistically significant difference was found between the total mean scores of the CCT-X pre-test and post-test for the experimental and control group students ($F_{(1,56)}$= 12.20, p<0.05). In terms of the measurement primary effect, no statistically significant difference was found between the CCT-X mean scores for students from the pre-test to post-test stages, regardless of gender ($F_{(1,56)}$= 2.26, p>0.05). When the partial eta squared values were analysed, it was found that the group factor explained 18% of the variance in the students' critical thinking skills scores; while the value of group and measurement factors together explained 9% of the variance in students' critical thinking skills scores (Table 10).

Discussion, Conclusion and Recommendations

According to the first results obtained in this study, it was found that the CCT-X pre-test scores for control group students, in which existing teaching methods were used, decreased when compared with the post-test scores, and there was no statistically significant difference (Table 4). Within this context, it can be said that the current teaching method applied to the control group students influenced the development of the critical thinking skills of seventh graders to a negative extent. The claim that the use of concept cartoons is effective in terms of facilitating an in-class discussion environment is supported by Seferoğlu and Akbıyık's (2006) view that the discussion technique is more relatively effective than simply lecturing in terms of giving students critical thinking skills. In addition, critical thinking skills depend on our ability to ask effective and smart questions, which delve deep into the subject and reveal its complex detail (Berman, 1985). In light of all these data, it can be said that, since the method applied in the control group is simply based on lecturing, students passively learn target behaviours as a result, without asking questions and not being able to express their thoughts freely. Furthermore, since there is no discussion environment and hardly any questions asked to cultivate the capacity to think, as preference is given to rote learning, the development of critical thinking skills is negatively influenced.

Another result of this study is the statistically significant difference between the CCT-X pre-test and post-test scores for the students in the experimental group, which favours the post-test (Table 7). This result shows that concept cartoons are effective in increasing the critical thinking skills of the experimental group students significantly. Wheeler and Collins (2003) examined the effects of concept cartoons used for preparing nursing students in their baccalaureate on students' critical thinking skills. Cho and Reich (2008) stated that cartoons can be used as teaching tools to encourage students to think critically and encourage them to participate in discussions. The findings of this and other studies reported in the literature support this claim from Cho and Reich (2008). In their study, Webb, Williams and Meiring (2008) concluded that concept cartoons were a promising teaching tool in terms of encouraging students to think and discuss. Köseoğlu (2009) concluded that using concept cartoons in social science lessons increased the critical thinking skills of seventh graders. The results of the study showed that concept cartoons were effective in helping the critical thinking skills of students.

Seferoğlu and Akbıyık (2006) emphasized the importance of open-ended questions and the method of discussion in developing students' critical thinking skills. This study, conducted with concept cartoons, was based on these two factors as a result of the nature of the method used. Students made connections with the concept cartoons displayed on the board by the teacher and given to them in paper form, then interpreted these connections and evaluated them. Thanks to this teaching method, students were guided to express their thoughts freely by being asked open-ended questions and were given the chance to discuss each other's views. Thus, it can be said that creating this kind of teaching environment is one of the reasons to explain the development of students' critical thinking skills. In addition, it can be stated that, since cartoons are colourful, witty and interesting, they encourage the active participation of students in lessons and thus play a contributive role in the development of critical thinking skills.

As a result of the study, it was found that there is a statistically significant difference in favour of the experimental group when comparing the critical thinking skill scores of the experimental group (in which concept cartoons were used) and the control group (in which existing teaching methods were used) in the seventh grade "Solar System and Beyond" unit (Table 9). Studies reporting parallel results with those for this study were also found (Miri, David & Uri, 2007; Vieira & Tenreiro-Vieira, 2016). Miri, David and Uri (2007) conducted a three-year-long longitudinal study, which aimed to examine whether teaching strategies, designed to develop higher order thinking skills in science education, developed the critical thinking skills of 11th graders. As a result of the study, they concluded that teaching applications, which aimed to develop higher order thinking skills, led to no statistically significant development on critical thinking factors and the tendency to cultivate the critical thinking subdimensions of students when compared with traditional teaching methods. Based on these results, it has been stated that, if teachers try to apply higher order thinking skills (for example, by dealing with real-life problems in class, encouraging open-ended in-class discussions and thought experiments) with purpose and in a persistent way, this

will create a better opportunity to develop critical thinking skills (Miri et al., 2007). Similarly, Vieira and Tenreiro-Vieira (2016) concluded that learning experiences within science learning and teaching, in which students aimed to be scientifically literate and were taught based on critical thinking, had a significant effect on the development of sixth graders' scientific literacy and critical thinking skills.

In light of the conclusions made by this study, the following recommendations can be made:

- Teachers have an important role to play in the development of critical thinking skills among students. Thus, practices such as pre-service teacher training programmes or in-service training can be increased in terms of quantity and quality.
- Studies can be conducted on the relationship between various methods and techniques used in the learning-teaching process for astronomy subjects and their impact on students' other thinking skills (problem-solving skills, reflective thinking skills, questioning skills).
- Various mixed research studies can be conducted to obtain more detailed information on students' critical thinking skills.
- In order to overcome the limitations of scales used in measuring critical thinking skills, studies should be conducted data diversity which data collection sets, such as observation, interview and document analysis, are used.
- This research consists of a three week teaching process. Longer-term longitudinal studies aimed at improving students' critical thinking skills can be conducted.
- In studies on providing students with and developing their critical thinking skills, multidimensional and extensive applications, blended with various teaching methods, rather than one teaching method, should be conducted, and their efficiency examined. In addition, the development of these applications can be observed by year on different sample groups, so that the efficiency of the technique applied can be assessed. In this way, the problems with the application can be identified and an effective method can be found to provide students with critical thinking skills.
- This research focuses on the development of critical thinking skills of students on subject-based topics in astronomy. Studies aiming at improving the critical thinking skills of students in faculty issues (physics, chemistry, biology, environmental education, socio-scientific subjects, etc.) can be done.

References

Akar, Ü. (2007). Öğretmen adaylarının bilimsel süreç becerileri ve eleştirel düşünme beceri düzeyleri arasında ilişki [The relationship between student teachers' scientific process skills and critical thinking]. Unpublished master's dissertation, Institute of Social Sciences, Afyonkocatepe University, Afyon.

Anonymous. (2016). Eleştirel düşünme becerileri okulu kitapçığı [Critical thinking skills school booklet]. Retrieved from http://cocukuniversitesi.aydin.edu.tr/belgeler/elestirel.pdf.

Aybek, B. (2007). Eleştirel düşünmenin öğretiminde öğretmenin rolü [The role of the teacher in the teaching of critical thinking]. *Bilim, Eğitim ve Düşünce Dergisi*, 7(2). Retrieved from http://www.universite-toplum.org/text.php3?id=322.

Bahrani, T. & Soltani, R. (2011). The pedagogical values of cartoons. *Research on Humanities and Social Sciences, 1*(4), 19-22.

Balım, A. G., İnel, D. & Evrekli, E. (2008). Fen öğretiminde kavram karikatürü kullanımının öğrencilerin akademik başarılarına ve sorgulayıcı öğrenme becerileri algılarına etkisi [The Effects the Using of Concept Cartoons in Science Education on Students' Academic Achievements and Enquiry Learning Skill Perceptions]. *İlköğretim Online*, 7(1), 188-202.

Berman, S. (1985). Thinking in context: teaching for open mindedness and critical understanding. In: Costa, A. L. (ed.). *Developing Minds: A Resource Book for Teaching Thinking*. Alexandria, VA: Association for Supervision and Curriculum Development.

Birisci P., Metin, M. & Karakas, M. (2010). Pre-service elementary teachers' views on concept cartoons: a sample from Turkey. *Middle East Journal of Scientific Research*, *5*(2), 91-97.

Browne, M. N. & Freeman, K. (2000). Distinguishing features of critical thinking classrooms. *Teaching in Higher Education*, *5*(3), 301-309.

Can, A. (2014). *SPSS İle Bilimsel Araştırma Sürecinde Nicel Veri Analizi [Quantitative Data Analysis in Scientific Research Process with SPSS]* (3rd Ed.). Ankara: Pegem Akademi Yayıncılık.

Chin, C. & Teou, L. Y. (2009) Using concept cartoons in formative assessment: scaffolding students' argumentation, *International Journal of Science Education, 31*(10), 1307-1332.

Cho, S. & Reich, G. A. (2008). New immigrants, new challenges: high school social studies teachers and English language learner instruction. *The Social Studies*, *99*(6), 235-242.

Dabell, J. (2008). Using concept cartoons. *Mathematics Teaching*, 209, 34-36.

Diakidoy, N. & Kendeou, P. (2001). Facilitating conceptual change in astronomy: a comparison of the effectiveness of two instructional approaches. *Learning and Instruction*, *11*, 1-20.

Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership*, *43*(2), 45-48.

Ennis, R. H. (1991). Critical thinking: a streamlined conception. *Teaching Philisopy, 14*(1), 5-23.

Field, A. (2009). *Discovering statistics using SPSS* (3rd Ed.). London: Sage Publication.

Fraenkel, J., Wallen, N. & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th Edition). McGrawHill: Boston.

Giancarlo, C. A. & Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *The Journal of General Education*, *50*(1), 29-55.

Ip, W. Y., Lee, D. T. F., Lee, I. F. K., Chau, J. P. C., Wootton, Y. S. Y. & Chang, A. M. (2000). Disposition towards critical thinking: a study of Chinese undergraduate nursing students. *Journal of Advanced Nursing*, *32*(1), 84-90.

Jiwaji, N. T. (2016). Recognition of astronomy as an essential discipline at all levels of education in Tanzania. *Rwanda Journal, 1*(1), 82-87.

Kabapınar, F. (2005). Effectiveness of teaching via concept cartoons from the point of view of constructivist approach. *Educational Sciences: Theory & Practice*, *5*(1), 135-146.

Keogh, B. & Naylor, S. (1999). Concept cartoons, teaching and learning in science: an evaluation. *International Journal of Science Education*, *21*(4), 431-446.

Keogh, B., Naylor, S. & Wilson, C. (1998). Concept cartoons: a new perspective on physics education. *Physics Education*, *33*(4), 219-224.

Keogh, B., Naylor, S., De Boo, M. & Feasey, R. (1999). Formative assessment using concept cartoons: initial teacher training in the UK. *Research in Science Education: Past. Present and Future*, 137-142.

Köseoğlu, M. (2009). *Sosyal bilgiler dersinde karikatür kullanımının eleştirel düşünmeye etkisi [The effect of using caricature in social studieson critical thinking].* Unpublished master's dissertation, Institute of Social Sciences, Gaziosmanpaşa University, Tokat.

Miri, B., David, B. C. & Uri, Z. (2007). Purposely teaching for the promotion of higher-order thinking skills: a case of critical thinking. *Research in Science Education*, *37*(4), 353-369.

MoNE [Ministry of National Education] (2010). *Ortaöğretim astronomi ve uzay bilimleri dersi öğretim programı [High school astronomy and space science curriculum].* Ankara.

MoNE [Ministry of National Education] (2013). *İlköğretim kurumları (ilkokullar ve ortaokullar) fen bilimleri dersi (3, 4, 5, 6, 7 ve 8. sınıflar) öğretim programı [Primary education institutions (primary and secondary schools) science cource (3, 4, 5, 6, 7 and 8 classes) curriculum].* Ankara.

Naylor, S. & Keogh, B. (2010). *Concept Cartoons in Science Education* (Second Edition). Sandbach: Millgate House.

Naylor, S. & Keogh, B. (2012). Concept cartoons: what have we learnt? Paper presented at the Fibonacci Project European Conference, Leicester, UK. Retrieved from http://www.millgatehouse.co.uk/research/concept cartoons-research-ongoing.

Norris, S. P. (1985). Synthesis of research on critical thinking. *Educational Leadership*, *42*(8), 40-45.

Özalp, I. (2006). *Karikatür Tekniğinin Fen ve Çevre Eğitiminde Kullanılabilirliği Üzerine Bir Araştırma [A study on implementing the technique of caricature in science and environmental education].* Unpublished master's dissertation, Institute for Graduate Studies in Science and Technology, Celal Bayar University, Manisa.

Özdemir, E. (2017). Humor in Elementary Science: Development and Evaluation of Comic Strips about Sound. *International Electronic Journal of Elementary Education, 9*(4), 837-850.

Percy, J. R. (2006). Teaching Astronomy: Why and How? *The Journal of the American Association of Variable Star Observers, 35*(1), 248-254.

Percy, J. R. (1998). Astronomy education: an international perspective. *Astrophysics and Space Science, 258*, 347-355

Schafersman, S. D. (1991). An introduction to critical thinking. Retrieved from http://www.freeinquiry.com/critical-thinking.html

Seferoğlu, S. S. & Akbıyık, C. (2006). Eleştirel düşünme ve öğretimi [Critical thinking and teaching]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *30*, 193-200.

Trumper, R. (2006). Teaching future teachers basic astronomy concepts – seasonal changes – at a time of reform in science education. *Journal of Research in Science Teaching*, 43(9), 879-906.

Tsai, P., Chen, S., Chang, H., & Chang, W. (2013). Effects of prompting critical reading of science news on seventh graders' cognitive achievement. *International Journal of Environmental & Science, 8*(1), 85-107.

Tümay, H. & Köseoğlu, F. (2011). Kimya öğretmen adaylarının argümantasyon odaklı öğretim konusunda anlayışlarının geliştirilmesi [The developing the understanding of chemistry teacher candidates about argumentation oriented teaching]. *Türk Fen Eğitimi Dergisi (TUSED)*, *8*(3), 105-119.

Uçar, S. & Demircioğlu, T. (2011). Changes in preservice teacher attitudes toward astronomy within a semester-long astronomy instruction and four-year-long teacher training program. *Journal of Science Education and Technology*, *20*, 65-73.

Vieira, R. M. & Tenreiro-Vieira, C. (2016). Fostering scientific literacy and critical thinking in elementary science education. *International Journal of Science and Math Education*, *14*(4), 659-680.

Webb, P., Williams, Y. & Meiring, L. (2008). Concept cartoons and writing frames: developing argumentation in South African science classrooms? *African Journal of Research in SMT Education*, *12*(1), 4-17.

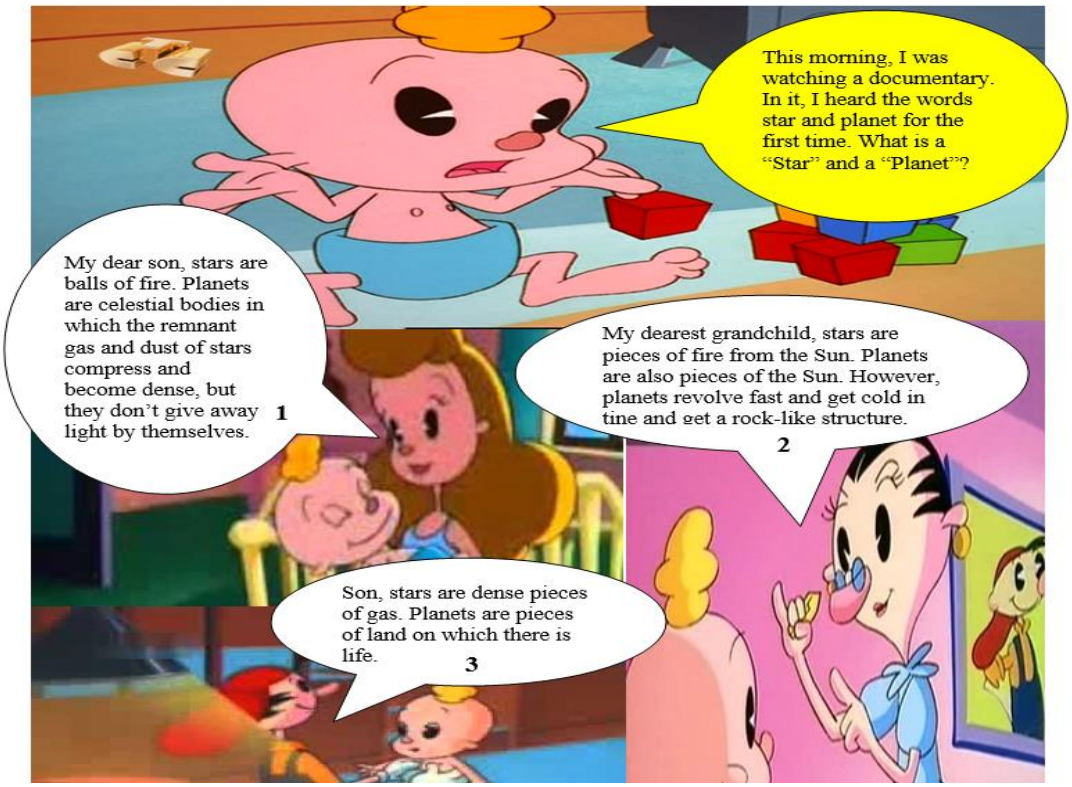
Wheeler, L. A. & Collins, S. K. R. (2003). The influence of concept mapping on critical thinking in baccalaureate nursing students. *Journal of Professional Nursing*, *19*(6), 339-346.

Zeilik, M., Schau, C., Mattern, N., Hall, S., Teague, K. W. & Bisard, W. (1997). Conceptual astronomy: a novel model for teaching postsecondary science courses. *American Journal of Physics*, *65*(10), 986-996.

Zousel, M. L., Rule, A. C. & Logan, S. R. (2013). Teaching primary grade students perfectionism through cartoons compared to bibliotherapy. *International Electronic Journal of Elementary Education, 5*(2), 199-218.

Appendix

Star and Planet Cartoon



The visuals in this concept cartoon were created by using the cartoon series named "The Why Why Family"

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

September 2017, Volume 10, Issue 2

Evaluation of Mathematical Game Design Skills of Pre-Service Classroom Teachers*

Pusat Pilten[a,**], Gülhiz Pilten[b], Ramazan Divrik[c], Fatma Divrik[d]

Received: 1 October 2017
Revised: 9 November 2017
Accepted: 17 November 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236120

Abstract

The purpose of the research is to evaluate the games prepared by pre-service classroom teachers within the scope of "Mathematics Teaching 1" and "Mathematics Teaching 2" courses, which are included in the undergraduate classroom teaching programs in Turkey, and to make predictions on the game design skills of pre-service teachers through them. The study has been conducted as a case study which is one of the qualitative research designs. The sampling group has been composed of 108 mathematical games designed by 386 undergraduate students, who were third year students at Necmettin Erbakan University, department of classroom teaching in 2015-2016 academic year, fall and spring semesters. Data collected through interviews and observations have been analysed with the help of descriptive analysis method. Results indicate that pre-service teachers failed at designing games that are satisfactory in terms of "Game Rules", "Goals and objectives", "Outcomes and feedback" and "Interaction". This can be taken as one of the main indicators of general failure in game designs. According to the other findings, game designs developed by pre-service teachers were partially or completely satisfactory in terms of "Conflict (Competition, challenge, and opposition)" and "Story"

Keywords: Mathematics education, pre-service classroom teachers, game design

Introduction

Games can be defined as activities that are done voluntarily and freely, provide a source of happiness, stimulate all developmental aspects of children and improve not only skills but also emotions (Razon, 1985). Additionally, the concept of game is taken from many different aspects and defined differently in the literature. Among these, Vygotsky (1967) states that game is a mechanism used by individuals to cope with the environment, and it forms the basis of development during pre-school period. According to him, abstract thought developed with games has a very important place in the development of individuals. According to Piaget (1962), every act of intelligence is defined by the balance between two opposite tendencies; assimilation and accommodation. In assimilation, individuals involve events, objects and cases within the existing ways of thinking that form the organizational mental structures. In accommodation, existing ways of thinking are combined with new aspects of external environment and re-organized. With the acts of intelligence, the individuals accommodate with the requirements of external reality and at the same time preserve their mental structures completely. On the contrary, game is defined with the superiority of assimilation over accommodation. Individuals involve events and objects within their existing mental structures. In other words, Piaget (1962) considers game as an "assimilation function". Accordingly, the studies on game in the literature mostly focus on cognitive development. This is duty the fact that game is defined as providing children with special and extra-ordinary environments, such as learning, enlightenment and developing meaning (Wood & Attfield, 2005).

In agreement with the conceptual framework provided above, and with the changes and improvements in the structures, rules and purposes of the games in line with the cognitive development of especially the pre-school and primary school children, games are today used as a method of teaching. Findings of many recent studies have showed that educational games are activities that have an important effect on the learning process (Shin, Sutherland, Norris, & Soloway, 2012; Virvou, Katsionis, & Manos, 2005). The common main finding of these studies is that games enable students do "a mental training". In this sense, students participating in games constantly encounter with decision-making cases during the game, and they use problem solving strategies for complex tasks involving interwoven sub-tasks (Johnson, 2005).

Gee (2003), defines these complex tasks and the cases students face within these tasks in four steps for video games: (1) engage/probe: gamers should engage in the existing situation and probe this; (2) Hypothesize: they

* A certain portion of this study has been presented in II. INES International Academic Research Congress in 18-21 October 2017, Antalya-Turkey.

[a,**]Corresponding author: Pusat Pilten, Khoja Akhmet Yassawi International Kazakh-Turkish University, Faculty of Human Sciences, Turkistan, Kazakhstan, E-mail: ppilten@hotmail.com
[b] Gülhiz Pilten, Khoja Akhmet Yassawi International Kazakh-Turkish University, Turkistan, Kazakhstan, E-mail: gulhizp@yahoo.com.tr
[c] PhD Candidate, Necmettin Erbakan University, Konya, Turkey. E-mail: rdivrik42@gmail.com
[d] Fatih Primary School, Afyonkarahisar, Turkey. E-mail: fk.divrik@mynet.com

should form a hypothesis by reflecting the experiences acquired through probing; (3) re-probe: they should re-probe using the hypothesis they formed mentally; (4) re-think: they should rethink the original hypothesis evaluating the feedbacks form the game environment.

Similarly, Garris, Ahlers, and Driskell (2002) state that educational games enable students encounter the processes of repetitive decision-making, acting and getting feedback. Additionally, the related literature includes many studies reporting that educational games are very important in developing decision-making, design, strategy, cooperation and problem solving skills of students, and accordingly they are valuable tools for learning and teaching processes (McFarlane, Sparrowhawk, & Heald, 2002; Cohen Group Report, 2011).

Taken the educational benefits mentioned above into consideration, games, which are of increasing and enriching learning quality, should be placed at the basis of the educational programs so that they can provide more benefits for the children. Educators should utilize these features of the games as they plan in-class activities (Morrison, 1998). Many studies on the subjects emphasize the importance of comprehending game world, imaginary world and real world of children for educators. Additionally, children can develop their skills with various knowledge and experiences by combining their experiences in these different worlds only with their educators' awareness and knowledge in the field their abilities to design effective games (Wood & Attfield, 2005; Ceglowsky, 1997).

Effective Educational Game Designs

The related literature includes many different models defining effective game design (Holland et al., 2003; Prensky, 2003; Dondlinger, 2007; Moreno-Ger et al., 2008; Shute & Ke, 2012; Whitton, 2012). These models present the elements of effective game design.

The models developed by Holland et al. (2003), Moreno-Ger et al. (2008) and Whitton (2012) define the dimensions of effective game design as: (1) Evaluation of game; (2) Cost-effectiveness; (3) Pedagogical effectiveness; (4) Gamification. Shuteve Ke (2012) proposed that core-elements of well-designed games must include: (1) Interactive problem solving; (2) Sensory Stimuli; (3) Uncertainty to gain engagement; (4) On-going feedback; (5) Control; (6) Adaptive Challenges; (7) Specific goals/rules.

Additionally, Dondlinger (2007) offered elements of effective video games as: (1) Edutainment; (2) Motivational attributes; (3) Goals and rules; (4) Narrative context; (5) Interactivity & multisensory. Van Staalduinenand de Freitas (2010) also listed 12 game elements from the literature and grouped them into four dimensions, which are shown in Figure 1. Some of the game elements appearing in the four dimensions are also mentioned in the core elements of a well-designed game and elements of an effective game (De Freitas & Jarvis, 2009; De Freitas & Oliver, 2006).

Four-Dimensional Framework	
Learner specifies	**Pedagogy**
Profile	Associative
Role	Cognitive
Competencies	Social-Situative
Representation	**Context**
Fidelity	Environment
Interactivity	Access to learning
Immersion	Supporting resources

Figure 1. Game elements grouped in a four dimensional framework

Similarly, the model of Prensky (2003) includes 6 key elements of game: (1) Rules: Rules define the borders of the game and provides us with some ways to reach our goals; (2) Goals and objective: goals and objectives provide gamers with a sense of duty and enable them play the game voluntarily and spend time and effort on it; (3) Outcomes and feedback: how the user proceeds towards the goal is notifies through feedbacks. Feedback is provided when something changes due to something user does within game; (4) Conflict: conflict, competition, challenge or opposition is the problem of the game that needs to be solves. The user can feel the senses of fear and excitement within the game as in real life, without encountering real danger, which motivates the user to continue playing and complete the game; (5) Interaction: is a social situation formed between gamers as they play the game; (6) Representation or story: is what the game is about. The story of the game can be presented at the beginning of the game or can be presented in more detail within the game.

Objective

A general evaluation of the information presented above shows that the effective use of games as an educational tool or teaching method especially focusing on cognitive development is very important and game design must be taken systematically for maximum educational benefit. To this end, presenting the game design competitions of primarily teachers and therefore the pre-service teachers is important. Most of the studies is the related literature focus on how learning is realized through games and the effects of the games on the learners (Garris, Ahlers, & Driskell, 2002; Gee, 2003; Kasvi, 2000; Pivec & Kearney, 2007). At this point, studies on educational game design are more important. Accordingly, the purpose of the present research is evaluating the games prepared by pre-service classroom teachers within the scope of "Mathematics Teaching 1 and Mathematics Teaching 2" courses, which are included in the undergraduate classroom teaching programs in Turkey, and making predictions on the game design skills of pre-service teachers through these.

In accordance with this purpose, the problem statement of the research is "What is the quality of the pre-service classroom teachers' game design skills within the framework of games they designed for the requirements of Mathematics Teaching 1 and 2 courses?" Answers to the following questions are sought in order to find an answer to this problem.

What is the quality of the games designed by pre-service teachers in terms of: (1) Rules; (2) Goals and objective; (3) Outcomes and feedback; (4) Conflict (Competition, challenge, opposition); (5) Interaction; (6) Representation or story?

Research Design

The present research on defining the quality of the game designs of pre-service classroom teachers is modelled as a case study. Case study can be defined as a research method that investigates phenomena in their own environment, with no distinct borders between the phenomena and their environment, based on "how" and "why" question, and enables the researchers to investigate in detail any phenomenon or event they cannot control (Yıldırım & Şimşek, 2005). What distinguishes case study from other methods is that it is based on 'how' and 'why', enables researchers to be included in a phenomenon or event they cannot control and investigate it in detail (Ekiz, 2009; Yıldırım & Şimşek, 2005). The reasons for using this model in the present research are, that the game designs of pre-service classroom teachers are accepted as a case with no distinct borders and that cannot be controlled by the researchers; and that this model enables researchers get involved in this case, investigate it in detail and find answers to "why-how" questions.

The present study is designed in accordance with "holistic single case study", which is one of the case study designs. Holistic single case studies involve one single analysis unit. It is used to confirm or reject a well-formulated hypothesis, to study unique or extreme cases, or for cases never studies or revealed before (Yıldırım & Şimşek, 2005). The single case studied in the present study is the game designs of pre-service teachers. The present study takes pre-service teachers as an analysis unit.

Participants

The present research is structured to analyse 108 mathematical games designed by 386 students, who were third year students at Necmettin Erbakan University, Department of Classroom Teaching in 2015-2016 academic year fall and spring semesters. The sample was formed in accordance with criterion sampling, which is a purposive sampling method. Purposive sampling enables the studying of cases, which are thought to have rich information (Patton, 1997). The main principle of criterion sampling is studying of the all cases meeting a series of pre-determined criteria. The criterion or criteria can be formed by the researchers of they can use a prepared criterion list (Yıldırım & Şimşek, 2005). The criterion for the selection of the pre-service teachers to participate in the present study was that they were informed of Turkish language teaching technique in various courses (Teaching Principles and Methods, Development and Learning, etc.) and they had practice in these subjects, they had completed their teaching practice.

Data Collection Tool

Data for the present research were collected via observation and interview techniques. Observations provide information on how the incidents actually happen (Çepni, 2007). The purpose of a simple observation is listening to what people say, monitoring what they do and sometimes asking some questions to provide clarity for some questions (Gillham, 2000). Observation technique is utilized to define an incident in detail (Yıldırım & Şimşek, 2005).

"Semi-structured game design observation form" used during the observations conducted for the present research was developed by researchers in accordance with the models provided in the related literature (Prensky, 2003). This form, includes the (1) Rules; (2) Goals and objective; (3) Outcomes and feedback; (4) Conflict (Competition, challenge, opposition); (5) Interaction; (6) Representation or story dimensions that define the structure of educational game according to the literature. In order to support observation findings, interviews are frequently utilized (Yıldırım & Şimşek, 2005). Therefore, unstructured interview questions were frequently used for data reliability.

Data Collection Process

The present research studies the games designed by 3rd year undergraduate classroom teaching students within the scope of "Mathematics Teaching 1 and Mathematics Teaching 2" courses. Observations and interviews made for this purpose have been conducted at the classrooms in the university building where the specified courses have been taught. At each session which the students presented their game designs, at least, two of the researchers have been provided to be observers. All of the students' applications related to the game designs have been recorded on video. Each student group's game design presentations lasted approximately 25 minutes. After the presentation of the game designs of the students, in order to be able to analyse the game design in more detail, questions have been asked by the researchers who follow the presentation. With this practice, it is aimed to support the observations realized by the interviews.

In the data collection process of the research the following steps were followed accordingly.

1. The participating students were informed about the theoretical framework related to the educational game concept.
2. Mathematics subjects related to the learning domains included in the primary school mathematics curriculum were distributed among students.
3. Students were provided with enough time for preparation (3 weeks).
4. Students were asked to organize the game play process for the games they designed in the classroom environment following the following steps.
   - 4.1 Introducing the game: This step involves the introduction of the game rules to the other students in the class. Researchers didn't interrupt at any point, but just observed.
   - 4.2 Application of the game: This step involves the playing of the game in the classroom

environment. Researchers also didn't interrupt at any point, but just observed during this step.

4.3 Evaluation of the game: After playing the game, students were asked to evaluate the game. During the evaluation, interviews were conducted by asking students some questions about the game design.

Data Analysis

Observations conducted for especially qualitative researches can be recorded with video, audio recorders or by taking notes (Yıldırım & Şimşek, 2005). Semi-structured observations conducted for the present research were video-recorded by the researchers. During observations, interviews were frequently utilized. While analysing the observation data, all researchers re-watched the game application, which were video-recorded by the researchers, and each researcher marked the related parts on the observation form to show their evaluations of the games and provided explanations. In other words, each researcher individually filled in the semi-structured observation form including the observation levels of criteria required by the game design dimension. After that, the researchers compared their forms together, discussed on each and made one form for each game when they had a consensus. In this process, the grading expressions of specialists for each participant has been compared with each other and reliability has been calculated using the formula of Miles and Huberman (1994) (Reliability=opinion association/ opinion association+opinion separation X100). With the help of the formula, it has been calculated that the correspondence among encoders in the range of 87.25 - 100%. Observation data obtained this way are presented in detail in findings part. Data collected through observations are coded and presented in tables, pre-service teachers' opinions are presented when necessary to support the observation findings and necessary explanations are provided under the tables.

Findings

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of game rules are presented in Table 1.

Table 1. *Evaluation of Game Designs in Terms of "Game Rules"*

Grade	Number of games	Rules							
		The rules of the game are far from being explanatory of the borders of the game and various ways of achieving the goal.		The rules of the game are barely explanatory of the borders of the game and various ways of achieving the goal.		The rules of the game are partially explanatory of the borders of the game and various ways of achieving the goal.		The rules of the game are completely explanatory of the borders of the game and various ways of achieving the goal.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	10	41.67	6	25	4	16.67	4	16.66
2	29	15	51.73	7	24.13	5	17.24	2	6.90
3	30	17	56.67	9	30	3	10	1	3.33
4	25	13	52	5	20	4	16	3	12

As presented in Table 1, almost half of the total of 108 games designed for all four grades of primary school are unsatisfactory in terms of rules (1st grade: 41.67%; 2nd grade: 51.73%; 3rd grade: 56.67%; 4th grade: 52%). According to the findings games designed by pre-service teachers were unsatisfactory in terms of defining the borders of the game and providing ways of achieving the goal. The rules set by the pre-service teachers mostly failed to provide an effective and uninterrupted game play. The reasons for this failure might be that pre-service teacher failed to predict the possible situations that can occur within the game and they didn't include any rules related to this situation. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"The game was supposed to be fun. But an unexpected situation occurred. That's why the game stopped."

"I couldn't predict …'s move in the game. I tried to come up with a new rule when … did that move but I guess my friends got a little bored meanwhile."

"The ones answering the difficult questions during the game were going to get higher scores. However, questions were not difficult but easy I think. That's why the game took to short. I should have thought of what to do when the game was over."

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of goals and objectives are presented in Table 2.

Table 2. *Evaluation of Game Designs in Terms of "Goals and Objectives"*

Grade	Number of games	Goals and objective							
		Learning objectives are either not existing or unclear. Goals don't create the sense of duty or voluntariness on the gamers.		Learning objectives were determined but it is not clear how the design elements reflect learning objectives. Goals create very little the sense of duty or voluntariness on the gamers.		Learning objectives were determined partially and some design elements reflect learning objectives. Goals create the sense of duty or voluntariness on the gamers.		Learning objectives are clear and design clearly reflects the learning objectives. Goals create a high sense of duty or voluntariness on the gamers.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	13	54.17	7	29.17	3	12.5	1	4.16
2	29	15	51.72	9	31.03	4	13.80	1	3.45
3	30	14	46.8	12	40	2	6.6	2	6.6
4	25	12	48	7	28	4	16	2	8

As presented in Table 2, almost half of the total of 108 games designed for all four grades of primary school are unsatisfactory in terms of "goals and objectives" (1st grade: 54.17%; 2nd grade: 51.72%; 3rd grade: 46.8%; 4th grade: 48%). According to the findings games designed by pre-service teachers were unsatisfactory in terms of defining the learning objectives and/or reflecting these on the game and creating a sense of duty and voluntariness. This seriously affected the participation in the game. The goals set for the games failed at providing motivation for the students to participate in the game. The reasons for this failure might be that pre-service teacher failed to set goals appropriate for the class level. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"It was fun when we played this game before. Yet, I think it didn't attract attention of my friends here. Nobody wanted to play. And we failed when we forced participation."

"The game started well. But later people got bored. When they got bored they played involuntarily. What could I do to prevent this? I think I should have new goals within the game while I designed the game."

"I think the game reached its objective. My topic was addition and we solved many operations. I think the only problem was that the students were university students. It would be much better if I did that with 4th graders."

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of outcomes and feedback are presented in Table 3.

Table 3. *Evaluation of Game Designs in Terms of "Outcomes and Feedback"*

Grade	Number of games	Outcomes and feedback							
		Learning experiences acquired in the designed game can't develop learning. Feedbacks aren't provided in time and correctly.		The game provides very few learning related experiences. Very few of the feedbacks are provided in time and correctly.		The game partially provides learning related experiences. Feedbacks are partially provided in time and correctly.		The game provides multiple learning related experiences. All of the feedbacks are provided in time and correctly.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	9	37.5	7	29.16	6	25	2	8.34
2	29	13	44.83	11	37.93	3	10.34	2	6.90
3	30	10	33.33	9	30	7	23.34	4	13.33
4	25	13	52	7	28	3	12	2	8

As presented in Table 3, most of the total of 108 games designed for all four grades of primary school is unsatisfactory in terms of "outcomes and feedbacks". According to the findings experiences acquired in the games designed by pre-service teachers were unsatisfactory in terms of developing learning experiences and providing feedbacks in time and correctly. This can be another source of the general failure of games. First of all, most of the games were away from learning objectives. Additionally, frequent and timeless repetition of feedbacks seriously affected the fluency of the games. The reason for frequent repetition of feedbacks can be that the games couldn't be explained correctly to the players. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"I thought I should have intervened when the game was misunderstood. It would be much better if I had provided some information at the beginning."

"The game got so far away from my design. So, I had to stop the game and tell them what they should do."

"I think the game failed at focusing on the math subject. My friends seemed to like running better."

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of "Conflict (Competition, challenge, opposition)" are presented in Table 4.

Table 4. *Evaluation of Game Designs in Terms of "Conflict"*

Grade	Number of games	Conflict (Competition, challenge, opposition):							
		Playing process has only a very few or no well-thought indicators providing players with conflict, competition, challenge and opposition.		Playing process has few well-thought indicators providing players with conflict, competition, challenge and opposition.		Playing process partially has well-thought indicators providing players with conflict, competition, challenge and opposition.		Playing process completely has well-thought indicators providing players with conflict, competition, challenge and opposition.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	5	20.83	4	16.67	9	37.5	6	25
2	29	5	17.24	1	3.45	13	44.83	10	34.48
3	30	4	13.33	3	10	12	40	11	36.67
4	25	3	12	2	8	11	44	9	36

As presented in Table 4, most of the total of 108 games designed for all four grades of primary school is partially or completely satisfactory in terms of "conflict". According to the findings experiences acquired in the games designed by pre-service teachers were satisfactory in terms of providing players with experiences of conflict, competition, challenge and opposition. This can be considered as the main factor providing playability in spite of the general failure resulting from rules, objectives, motivation and feedbacks. Pre-service teachers could manage to include especially the competition factor. Those who didn't stated that they didn't include the competition factor as they thought it is not suitable for primary school children. Additionally, some, who did, failed to adjust the level of competition. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"I didn't want competition in my game design because I didn't want the children to play competitively."

"I thought that if my game design included competition, more people would participate, but I think I overdid it. Element related to course content should have been more prominent."

"I think this kind of competitive games is not appropriate for children. We shouldn't raise humans who just enjoy winning. Human beings should also enjoy the activity they are in."

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of "Interaction" are presented in Table 5.

Table 5. *Evaluation of Game Designs in Terms of "Interaction"*

Grade	Number of games	Interaction							
		Game design involves the active participation of only one or two group members. It doesn't provide social interaction.		Game design involves the active participation of some group members. It provides partial social interaction.		Game design involves the active participation of most group members. It mostly provides social interaction.		Game design involves the active participation of all group members. It provides a high social interaction.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	14	58.33	4	16.67	3	12.5	3	12.5
2	29	18	62.07	7	24.13	2	6.90	2	6.90
3	30	17	56.67	8	26.67	3	10	2	6.66
4	25	15	60	5	20	3	12	2	8

As presented in Table 5, most of the total of 108 games designed for all four grades of primary school is unsatisfactory in terms of "interaction". According to the findings, games designed by pre-service teachers were unsatisfactory in terms of providing social interaction between group members. This failure can be resulted from lack of motivation, or the problem in rules. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"My game design provided opportunities for the participation of all class. But most of my friends didn't participate. I think failed to attract their interest."

"One member of each team would come to the board. After that, they would replace with other members of their teams. I think the ones on their seats got bored. When nobody wanted to come to the board, I completed the game with the ones who did."

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of "Story" are presented in Table 6.

Table 6. *Evaluation of Game Designs in Terms of "Story"*

Grade	Number of games	Story							
		The story of the game included very few or no features related to the development levels of students.		The story of the game included few features related to the development levels of students.		The story of the game partially included features related to the development levels of students.		The story of the game included many features related to the development levels of students.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	2	8.33	5	20.83	11	45.84	6	25
2	29	5	17.24	6	20.69	10	34.48	8	27.59
3	30	5	16.67	9	30	11	36.67	5	16.66
4	25	3	12	4	16	12	48	6	24

As presented in Table 6, most of the total of 108 games designed for all four grades of primary school is partially or completely satisfactory in terms of "story". According to the findings story element of the games designed by pre-service teachers were satisfactory in terms of including features related to the development level of students. Pre-service teachers paid attention to the stories of their games and considered this as the most important element of their designs. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"I think story is the most important element of the game. Stories should be appropriate for the ages of the children. In other words, children should live the story."

"The story setup of what is done in the game should be good. This provides more participation. They love the game more."

"The topics should be popular among children. Characters in the topics are also important. Children today might not love the heroes of our childhood. We should provide an environment, which children can love with contemporary characters."

The findings of the analyses conducted to evaluate the games designed by pre-service teachers in terms of "Representation" are presented in Table 7.

Table 7. *Evaluation of Game Designs in Terms of "Representation"*

Grade	Number of games	Representation							
		Characters, environment, roles, rules and other game elements fail to represent learning objectives.		Characters, environment, roles, rules and other game elements barely represent learning objectives.		Characters, environment, roles, rules and other game elements partially represent learning objectives.		Characters, environment, roles, rules and other game elements completely represent learning objectives.	
		f	%	*f*	%	*f*	%	*f*	%
1	24	9	37.5	7	29.17	5	20.83	3	12.5
2	29	14	48.28	8	27.59	6	20.69	1	3.44
3	30	18	60	8	26.67	2	6.67	2	6.66
4	25	11	44	6	24	5	20	3	12

As presented in Table 6, most of the total of 108 games designed for all four grades of primary school is unsatisfactory in terms of "representation". According to the findings, games designed by pre-service teachers were unsatisfactory in terms of representing the learning objectives by their characters, environment, roles, rules and other game elements. This shows that game designs failed to reach its educational objectives. Some of the statements supporting this finding obtained through interviews conducted during the evaluation of the game are as follows:

"The game was designed for 3rd grade geometry learning domain. But I think it included few math related elements."

"I thought students would understand the subject better after playing the game, but it didn't go as I expected. It was more fun than educational."

"I think my game works more as a practice or review than teaching new topics or raising awareness on something new."

Results and Discussion

The first finding of the research was that, pre-service teachers failed at designing games that are satisfactory in terms of rules. This can be taken as one of the main indicators of general failure in game designs. According to Dondlinger (2007) game rules are one of the elements of effective design. Many studies in the related literature emphasize the importance of rules in an effective design (Waraich, 2004; Zagal, Nussbaum, & Rosas, 2000). Especially in experimental game models formed to gather game design elements in accordance with educational theories underline that games provide a meaningful environment for problem-based learning and games should provide students with environments to discover new ideas and rules instead memorizing things (Kiili, 2005). The finding related to pre-service teachers' failing at defining game rules correctly indicates that the games designed by pre-service teachers will fail to reach at the educational objectives mentioned above.

Another finding of the present research is that, games designed by pre-service teachers were unsatisfactory in terms of "goals and objectives" as well. This negatively affected the continuity of the game and prevented obtaining the expected educational benefits from the games. Swartout and van Lent (2003), state that goals of different levels motivate learners to continue the game. This is considered as the main problem of the games studied within the scope of the present research. Pre-service teachers set simple goals, which they thought as appropriate for the developmental levels of primary school children. Yet, game levels changing in accordance with the levels of each learner and their success within the game increase the functionality of games (Gee, 2003). Appropriateness of the goals affects gamers' levels of commitment to the game. Another factor affecting commitment levels is motivation. Waraich (2004) reports that context should be of appropriate quality for meaningful learning and motivation is necessary for learning performance. The positive effects of correctly set goals on motivation have been proven by many researchers before (Amory, Naicker, Vincent, & Adams, 1999; Denis & Jouvelot, 2005; Jennings, 2001).

Third finding of the present research was that game designs of pre-service teachers were also unsatisfactory in terms of "outcomes and feedback". Because pre-service teachers couldn't explain the rules and goals of their games correctly, they had to stop the game and give feedback for correcting the acts that were against rules. Providing a certain level of interaction and feedback is a factor enabling players to carry on with the next goal and finally completely the game in victory. Yet, too much feedback within the game affects the players negatively (Song & Zhang, 2008). The related literature takes feedback in game design in two dimensions; player feedback and system feedback (Van Staalduinen & De Freitas, 2010). Appropriated briefing and system feedback will be given by the game according to the users' achievements to assist the students learning. The students learn from the system feedback and improve (player feedback). The players' feedbacks will influence the learning contents (maintain the current content, move to a more advance topic or go deeper into the same topic) and this will influence learning objectives and player goals. It was observed that game designs studied within the scope of the present research mostly used system feedback while ignoring player feedback. This is an indicator of failure of designs in terms of outcomes and feedback.

According to the findings, game designs developed by pre-service teachers were partially or completely satisfactory in terms of "Conflict (Competition, challenge, opposition)". Defining the game style in game design is very important in achieving the objectives of playing an educational game and forming the expected behaviours (Amory, 2001; Amory, Naicker, Vincent, & Adams, 1999; Ju & Wagner, 1997; Moreno Ger, Burgos, Martínez-Ortiz, Sierra, & Fernández-Manjón, 2008). The style that requires the cooperation between players to achieve the goals and the game style that requires competition between players with a winner and looser at the end are expected to be different. The findings of the present research show that pre-service teachers mostly preferred the competition-based games in their designs and they successfully adapted competition element in their games.

Another finding of the present research was that, game designs developed by pre-service teachers were unsatisfactory in terms of "interaction". This result in failure in providing an intense interaction environment required for an effective learning environment, and therefore failure in achieving the learning objectives of the designed games (Pivec & Kearney, 2007). Additionally, pre-service teachers' failure in reflecting the interaction factor on their educational game designs can prevent their designs to be accepted as an educational game (Denis & Jouvelot, 2005).

According to the findings of the present research, games designed by pre-service teachers were partially or completely satisfactory in terms of "story". Pre-service teachers commonly stated that choosing contemporary stories and characters and effective environment were important and additionally there shouldn't be disconnections within story of the game. The more realistic the stories of the games are, the more connected can be the players with the game. In games, players can see themselves as heroes taking active parts in tasks (Song & Zhang, 2008). This is also considered as one of the most important elements providing players with motivation (Dickey, 2005, 2006; Fisch, 2005; Waraich, 2004). Dickey (2006), stated that story context that stimulates the concepts of competition, imagination and curiosity increases internal motivation and participation related strategies can also increase motivation by covering options of role-play, story jumps, challenges and interaction. Despite the games designed within the scope of the present research lacked motivation due to unsatisfactory goal setting and feedbacks, partial success

resulting from motivation in game designs can be associated with the satisfactory story setting.

The final finding of the present research was that game designs developed by pre-service teachers were also unsatisfactory in terms of "representation". This finding indicates that targeted educational situations weren't represented within the game. Game factor in pre-service teachers' game designs got ahead of educational objectives. However, the balance between educational objectives and game play should be well adjusted in educational game design (Kiili, 2005). Overall evaluation of all game design elements implies that pre-service teachers failed at achieving planned efficacy from educational games in their mathematical game designs.

Following suggestions were developed in accordance with the findings presented above: (1) Pre-service teachers should be provided with theoretical and practical knowledge on game design theories; (2) Pre-service teachers should be provided with opportunities to develop game designs in other teaching courses included in undergraduate classroom teaching program (Language Teaching, Social Sciences Teaching, Science Teaching); (3) Pre-service teachers should be provided with opportunities to apply their game designs on the appropriate audience in teaching practice courses.

References

Amory, A. (2001). Building an educational adventure game: Theory, design and lessons. *Journal of Interactive Learning Research*, *129*(2/3), 249-264.

Amory, A., Naicker, K., Vincent, J., & Adams, C. (1999). The use of computer games as an educational tool: Identification of appropriate game types and game elements. *British Journal of Educational Technology*, *30*(4), 311-321.

Ceglowsky, D. (1997) Understanding and Building upon Children's Perceptions of Play Activities in Early Childhood Programs. *Early Childhood Education Journal*, *25*(2), 107-112.

Cohen Group Report. (2011). *Young children, Apps & iPad*. Retrieved from http://mcgrc.com/wp-content/uploads/2012/06/ipad-study-cover-page-report-mcg-info_newonline.pdf. last visit 11/08/17.

Çepni, S. (2007). *Araştırma ve Proje Çalışmalarına Giriş*. Trabzon: Celepler Matbaacılık.

De Freitas, S., & Jarvis, S. (2009). *Towards a development approach for serious games. Games-based learning advancements for multi-sensory human computer interfaces: Techniques and effective practices*. IGI Global. Hershey, PA.

De Freitas, S., & Oliver, M. (2006). How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? *Computers & Education*, *46*(3), 249-264.

Denis, G., & Jouvelot, P. (2005). *Motivation-driven educational game design: applying best practices to music education*. Paper presented at the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology, Valencia, Spain.

Dickey, M.D. (2005). Three-dimensional virtual worlds and distance learning: Two case studies of Active Worlds as a medium for distance education. *British Journal of Educational Technology*, *36*(3), 439-451.

Dickey, M.D. (2006). *"Ninja Looting" for instructional design: The design challenges of creating a gamebased learning environment*. Paper presented at the ACM SIGGRAPH 2006 conference, Boston.

Dondlinger, M.J. (2007). Educational video game design: A review of the literature. *Journal of Applied Educational Technology*, *4*(1), 21-31.

Ekiz, D. (2003). *Eğitimde Araştırma Yöntem ve Metotlarına Giriş: Nitel, Nicel ve Eleştirel Kuram Metodolojileri*. Ankara: Anı Yayıncılık.

Fisch, S.M. (2005). *Making educational computer games "educational".* Paper presented at the 2005 Conference on Interaction Design and Children, Boulder, CO.

Garris, R., Ahlers, R., & Driskell, J. (2002). Games, motivation and learning: a research and practice model. *Simulation and Gaming*, *33*(4), 441–467.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York, NY: Palgrave Macmillan.

Gilham, B., (2000). *Case Study Research Methods*, London: Continuum.

Holland, W., Jenkins, H., Squire, K., Miller, Z. N., O'Driscoll, A., Shresthova, S., Todd, K. (2003). Theory by design. *Video game theory reader*, 25-46.

Jennings, M. (2001). *Best practices in corporate training and the role of aesthetics: Interviews with eight experts*. Paper presented at the 2001 ACM SIGCPR Conference on Computer Personnel Research, SanDiego, CA.

Johnson, S. (2005). *Everything bad is good for you: How today's popular culture is actually making us smarter*. London: Allen Lane.

Ju, E., & Wagner, C. (1997). Personal computer adventure games: their structure, principles, and applicability for training. *SIGMIS Database*, *28*(2), 78-92

Kasvi, J. (2000). Not just fun and games. -Internet games as a training medium. *Cosiga- Learning with Computerised Simulation Games*. pp. 23-34.

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *Internet and Higher Education*, 8, 18-20.

McFarlane, A., Sparrowhawk, A., & Heald, Y. (2002). *Report on the Educational Use of Games: an Exploration by TEEM of the Contribution Which Games Can Make to the Education Process*. Cambridge: Teem

Moreno-Ger, P., Burgos, D., Martínez-Ortiz, I., Sierra, J. L., & Fernández Manjón, B. (2008). Educational game

design for online education. *Computers in Human Behavior*, *24*(6), 2530-2540.

Morrison, G.S. (1998). *Early Childhood Education Today*. Upper Saddle River, New Jersey

Patton, M., (1997). *Utilization-Focused Evaluation*, 3rd Edition. Thousand Oaks: Sage Publications

Piaget, J. (1962). *Play, Dreams, and Imitation in Childhood*. New York: Norton.

Pivec, M.,& Kearney, P. (2007). Games for learning and learning from Games. *Organizacija*, *40*(6), 419---423.

Prensky, M. (2003). Digital game–based learning. *Computers in Entertainment (CIE), 1*(1), 21-21

Razon, N. (1985). *Okul Öncesi Eğitimde Oyunun, Oyunda Yetişkinin İşlevi*. Ya-Pa Okulöncesi Eğitimi ve Yaygınlaştırılması Semineri. İstanbul.

Shin, N., Sutherland, L. M., Norris, C. A., & Soloway, E. (2012). Effects of game technology on elementary student learning in mathematics. *British Journal of Educational Technology*, *43*, 540–560.

Shute, V., & Ke, F. (2012). *Games, learning, and assessment*. Assessment in Game-Based Learning. Springer US. p. 43-58.

Song, M., & S. Zhang (2008). *EFM: A Model for Educational Game Design, in Lecture Notes in Computer Science*, S. Link, Editor. 2008, Springer US. p. 509-517.

Swartout, W., & van Lent, M. (2003). Making a game of system design. *Communications of the ACM*, *46*(7), 32-\--39.

Virvou, M., Katsionis, G., & Manos, K. (2005). Combining software games with education: evaluation of its educational effectiveness. *Educational Technology and Society*, *8*(2), 54–65.

Van Staalduinen, J. P. & de Freitas, S. (2010) *A Game-Based Learning Framework: Linking Game Design and Learning Outcomes*. In Learning to Play: Exploring the Future of Education with Video Games. Khine, MyintSwe (ed.). Peter Lang Publishers New York.

Vygotsky L. (1967). Play and its Role in the Mental Development of the Child. *Soviet Psychology*, *5*, 6-18.

Waraich, A. (2004). *Using narrative as a motivating device to teach binary arithmetic and logic gates*. Paper presented at the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education, Leeds, United Kingdom.

Whitton, N. (2012). The place of game-based learning in an age of austerity. *Electronic Journal of e-Learning*, *10*(2), 249-256.

Wood, E., & Attfield, J. (2005). *Play, learning and the early childhood curriculum* (2nd ed), London: Paul Chapman.

Yıldırım, A., & Şimşek, H. (2005). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri*. Ankara: Seçkin Yayınları.

Zagal, J.P., Nussbaum, M., & Rosas, R. (2000). A model to support the design of multiplayer games. *Presence*, *9*(5), 448-462.

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effects of Teaching Informative Text through Processual Model on Reading Comprehsion Skills*

Süleyman Erkam Sulak [a,**], Firdevs Güneş [b]

Received: 6 Sept. 2017
Revised: 18 October 2017
Accepted: 25 October 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236121

Abstract

This study was carried out in order to search the effects of teaching informative text structures through processual model on the reading comprehension skills of 4th grade students. The research was designed in accordance with experimental model with pre-test-post-test control groups. The study group of the study consisted of 62 fourth grade students who receive their education in 2012-2013 school year in a state school located in the province of Konya, Turkey. Teaching of the informative texts was carried out for 10 weeks based on the processual model in the experimental group, and based on the curriculum of Turkish course in the control group. As data collection tools, Reading Comprehension Test and Awareness Test of Informative Text Structures were used in the research. Mean, standard deviation, unpaired t-test, and one-way analysis of variance were used in the analyses of the study data. Results of the study have revealed that there are significant differences between the reading comprehension levels and awareness of informative text structures on behalf of the experimental group that learned informative texts through processual model.

Keywords: Reading comprehension, informative text, text structures, processual model

Introduction

Increasing the effectiveness of the reading comprehension skills is only possible with the use of various strategies. Using different strategies enables students and teachers to teach and learn texts, and to understand the topic (Dönmez & Yazıcı, 2006). Good readers have advanced strategies that they use to comprehend what they read (Vacca & Vacca, 2005).

The demonstration of comprehension strategies by teachers and the teaching of these strategies as models or guides will enable students to learn and use them effectively. Students who successfully learn these strategies will be readers who are independent from their teachers (National Reading Panel [NRP], 2000). At the end of the provided reading trainings, it should be aimed to develop independent readers who can actively use comprehension strategies.

Recently, reading researchers have been especially interested in comprehension trainings and developed various comprehension models. İnteractive and processual comprehension models, which are two of the comprehension models developed within the frame of constructivist approach, are commonly known and used (Güneş, 2007).

Processual model is a comprehension model that was developed by Kintsch and Van Dijk (1978). In the related literature, the processual model has names such as situation, construction-integration and comprehension model. According to processual model, in order to understand a text, the information given in the text should be reconstructed in a systematic way. It works in accordance with the principle of creating mental representations in the process of understanding texts. In this model, comprehension processes including the process and mental construction of the information given in the text are explained. At the level of micro structures and macro structures, the information given in the text are chosen, ranked, and organized in accordance with the prior knowledge of the reader, structure of the texts, and the rank of importance. At the level of micro structures, construction of the meaning of a word, a clause, or a sentence is conducted. On the other hand, at the level of macro structures, the meaning of the paragraphs, significant parts, or the whole of the text is conducted. At the same time, organizational structure of the text (description, sequence, cause and effects etc.) is also a macro structure that makes the text easy to understand. Comprehension of the text occurs as a result of the association of the micro and macro structures and prior knowledge of the reader. As these processes are repeated, automatic comprehension skills of the students will improve. With its different aspect, this model is not a model in which only meaning units are generated. It is a model that includes components such as visual images and emotions along with personal experiences (Kintsch & Dijk, 1978; Kintsch, 2002; Kintsch & Kintsch, 2005).

* This study is a part of Süleyman Erkam Sulak's doctoral dissertation: "The Effect of Procedural Model-Based Informative Text Instruction on Reading Comprehension Skills" supervised by Firdevs Gunes.

[a,**] Corresponding author: Suleyman Erkam Sulak, Ordu University, Faculty of Education, Ordu, Turkey. Phone: +905398779411, E-mail: erkamsulak@gmail.com

[b] Ankara University, Faculty of Education, Ankara, Turkey. E-mail : firdevs.gunes@gmail.com

Stages of Processual Model

Processual model assumes that as a mental skill, comprehension occurs as a result of the mutual interaction of various stages. The reader can reach the meaning of a text as a result of some stages. Cha and Swaffar (1998) stated the stages of processual model as follows:

Stage 1: First of all, the reader determines what is told in the text. At this stage, the topic of the text creates the focal point of the reader.

Stage 2: The reader determines how the text is organized based on the structural signs and the logical relations between the given information. In other words, s/he explains in which structure the text was written.

Stage 3: At this stage, the reader analyses the organization of the text. In order to do that, s/he focuses on the words, sentences, details, and supporting ideas in the text. S/he identifies the relations between ideas by being aware of the structure of the text, and s/he creates supporting ideas. S/he presents the created supporting ideas as a list. S/he reaches the main idea by associating the supporting ideas provided as a list with each other. In order to reach the main idea of the text, the reader should create the links between the ideas in the text again.

Stage 4: At this stage, the reader unifies her/his own perspective and the main idea of the text in order to make inferences. By this means, s/he creates her/his own meaning of the text.

Processual model is an effective model that can be used in the teaching of informative texts, and hence in the development of the comprehension skills (Weaver & Kintsch, 1996). For this reason, in this study, processual model was used in the teaching of informative texts.

Teaching Informative Texts

With the Turkish course curriculum dated 2005, the texts were begun to be taught with a thematic approach, and students were begun to be presented various text structures in the genres of narrative, informative, and poetry. This situation indicates that teachers should also draw the attention to the structural differences of the texts in the process of teaching (Akyol, 2007).

Teaching of the informative texts is carried out in two stages;

1. Introducing different informative text structures,
2. Using basic comprehension strategies for each structure (Simonsen, 2004).

Structure of a text reflects the organization of ideas in the text and the relations between these ideas (Armbruster, 2004). In order to improve the comprehension skills of the students, students should be clearly taught how to recognize and use text structures (Dreher & Gray, 2009; Dymock, 2005; Minskoff, 2005).

While narrative texts, which are more widely known and loved by the students, are included in the first years of elementary school, the number of informative texts increases in the following years. In the process of transition from narrative texts to informative texts, students generally have difficulty in reading and comprehending these texts since their understanding of informative text structures is not developed (Vacca & Vacca, 2005).

While new vocabulary and terms are more highly given in informative texts, direct personal experiences are less mentioned (Hall, Sabey & Mcclellan, 2005). At the same time, they are harder to understand compared to narrative texts since they have different structures (Williams, 2005).

Although there is not a mutual classification, researchers generally classify informative text structures in five groups:

1. Description
2. Sequence
3. Cause and effect
4. Compare and contrast
5. Problem solving (Gunning, 2005; Meyer,1985, as cited in Moss, 2004; Rozmiarek, 2006; Simonsen 2004; Temple et al. 2005; Vacca & Vacca, 2005; Williams, 2005).

Informative text structures should be taught separately in company with the appropriate strategies (Dymock, 2005; Gunning, 2005) because the quality of a structure cannot be transferred to other structures (Williams, 2005). For example, in order to understand a text written in the type of "sequence" text structure of a text written in the type of "comparison" cannot be used.

In the teaching of informative texts, identifying of the structure of a text is only the first stage; in the following stage, readers should choose important ideas given in the text and associate them with each other (Gunning, 2005).

Informative texts include topic, main idea, and supporting ideas (Akyol, 1999). Ideas in the text are organized in text structures. A reader who uses the organizational structure of the text can more easily determine the main idea and supporting ideas (Gunning, 2005).

In the teaching of informative texts through processual model, the necessary information about how to apply processes of selection, sequence, and organizing to the paragraphs of the text should be given. It should be explained how the meaning of the text can be generated at the level of micro and macro structures (Kintsch & Van Dijk, 1978). Also, in the teaching of these types of texts, in order to present the information in a systematic way, graphic organizers should be used, and clue words that are frequently used in the structure of each text should be introduced (Gunning, 2005; Vacca & Vacca, 2005).

In the first years of elementary school, most teachers do not attach much importance to informative texts. One of the reasons of this situation is that teachers assume that students can understand when they read informative texts (Read, Reutzel & Fawson, 2008). This situation causes a significant decrease in the reading success of the students after the third grade of the elementary school.

Thus, in the following years, students find it difficult to keep up with their increasing demand to understand. The effective teaching of informative texts especially at the level of 4th grade has a critical importance in students' understanding of informative texts that they will come across in the upcoming periods (Hall, Sabey & McClellan, 2005). For this reason, informative texts should be taught from the very first years of the educational process with appropriate models, methods, and strategies.

In Turkey, there are a limited number of studies carried out about the teaching of informative texts. Some of the conducted researches focus on the summarizing skills in informative texts (Çakır, 1995; Çıkrıkçı, 2004; Görgen, 1997; Keçik, 1993). And some researches were carried out towards university students. For example, in her study, Kuzu (2003) searched the effects of the reading training based on transactional model on the junior college students' understanding level regarding to informative texts. However, at the level of elementary school, there is not any study in which informative texts are taught through a model and its effects on the students' comprehension levels are tested. For this reason, a study about this subject was needed.

The aim of this study is to identify the effects of teaching informative text structures through processual model on the reading comprehension skills of 4th grade students. In order to reach this aim, answers of the following questions were looked for:

1. Is there a significant difference between *"the informative text structure awareness"* of the students in the control group and of the students in the experimental group who have learned informative text types (description, sequence, cause and effect, and problem solving) through processual model?
2. Is there a significant difference between *"comprehension levels"* of the students in the control group and of the students in the experimental group who have learned informative text types (description, sequence, comparison, cause and effect, and problem solving) through processual model?

Methodology

Research Model

The research was designed in accordance with quasi-experimental model with pre-test-post-test control groups. In the quasi-experimental research model, participants are not objectively appointed to experimental and control groups. The researcher uses the available groups (Creswell, 2003). For this reason, quasi-experimental model is an efficient model that can be used especially in educational studies.

Participant

The study group of this research consisted of 62 fourth grade students who receive their education in 2012-2013 school year in a state school, which is considered to be in middle socio-economic class and located in the province of Konya, Turkey.

In order to determine the experimental and control groups, reading comprehension test in the type of *description* was applied to the 4th graders of the experimental school. It was tested through one-way analysis of variance (ANOVA) whether there was a significant difference between the scores (gotten out of 100) of the classes or not; and the results are given in Table 1.

Table 1. *ANOVA results of reading comprehension pre-test scores of the students in the experimental school*

Class	*n*	*M*	SD	*f*	*p*
4A	32	28.28	12.61		
4B	30	25	11.96		
4C	30	29.17	20.26	0.967	0.427*
4D	30	28.67	10.90		
4E	30	23.33	14.28		

*$p>0.05$

When Table 1 is analysed, it is observed that raw scores vary between 23.33 and 29.17. However, there was not a statistically significant difference between the scores gotten from the test by the classes ($p>0.05$). Therefore, experimental groups and control groups were randomly chosen from the classes of which means were close to each other. Within this context, between the matched classes, 4-D class was chosen as the experimental group and 4-A class was chosen as the control group.

Measures

Reading Comprehension Test: The test was developed by the researcher through analysing the relevant literature and obtaining expert opinions. In order to prepare the reading comprehension test, first of all, reading comprehension achievements for the 4th graders who were in the 2005 Turkish course curriculum were identified. A 65-question reading comprehension test including open-ended and multiple-choice questions, which were appropriate for the selected achievements, was prepared. The prepared reading comprehension test was applied to 115 students who have received education in the 4th grade of a state school.

In the item analysis carried out at the end of the practice, the items of which discrimination index were under .20 and the items in which there was not any difference between independent samples t-test and 27% slice of the subgroups and supergroups were not considered as distinctive, and 15 questions were omitted from the test. As a result, a 50-question *Reading Comprehension Test* was acquired. Reading comprehension test consisted of five parts (description, sequence, comparison, cause and effect, and problem solving). In each comprehension test, there were 10 questions as being 5 multiple-choice questions and 5 open-ended questions.

Cronbach's alpha coefficients of the reading comprehension test were calculated as follows: 0.75 for comprehension test of the description type, 0.75 for comprehension test of the sequence type, 0.72 for comprehension test of the comparison type, 0.76 for comprehension test of the cause and effect type, and 0.81 for comprehension test of the problem solving type. Reliability coefficient of the general reading comprehension test (50-items) was 0.80.

During the scoring of test items, for the multiple-choice questions, wrong answers were calculated as 0 point and correct answers were calculated as 1 point; and for the open-ended questions, wrong answers were calculated as 0 point, incomplete answers were calculated as 1 point, and correct answers were calculated as 2 points.

Awareness Test of Informative Text Structures: The test was developed by the researcher through analysing the relevant literature and obtaining expert opinions. In order to assess the students' knowledge about informative texts, short paragraphs from Turkish textbooks, which presented informative text structures, were selected. Questions about the structures of these paragraphs were asked in multiple-question form. Also, clue words that were frequently used in informative text structures and questions about the structural qualities of the texts were used in the test. As a result of the expert opinions, the awareness test of informative text structures consisting of 15 multiple-choice questions was acquired. For item analysis and score reliability, the test was applied to 120 students who have received education in the 4th grade of a state school.

When item discrimination coefficients of the test items were analysed, it was observed to be higher than 0.20. Therefore, no item was omitted from the test. The test's Cronbach's alpha coefficient was calculated as 0.76.

Process

The lessons in the experimental group were carried out by the researcher while the lessons in the control group were carried out by the class teacher. Studies were carried out for 10 weeks as being 2 weeks for the awareness program about the conducted method and 8 weeks for the practice process.

In the experimental group, the teaching of the informative texts was carried out in accordance with processual model. Within the scope of the stages stated by Cha and Swaffar (1998), various activities and strategies before, during, and after the reading of informative texts were used. These activities and strategies were prepared through using the studies of several researchers (Duke & Pearson, 2002; Hall, Sabey & McClellan, 2005; Hoffman, 2010; Gunning, 2005; Güneş, 2007; Karatay, 2011; NRP, 2000; Schirmer, 2010; Temple et al., 2005; Tompkins, 2006; Vacca & Vacca, 2005).

Before the reading, activities and strategies of creating the aim of reading, activating the prior knowledge, checking the title and text, estimating, brain storming, studying with the vocabulary, and KWL were used.

During the reading, reading activities, activities of underlining the important words and statements, finding clue words, and questioning and answering were used.

After the reading, activities of retelling, identifying the topic, finding the supporting ideas, finding the main topic, identifying the aim of the writer, graphic organizers, and summarizing were used.

Each text structure was separately taught in the experimental group. First of all, the training was started with the text structure of the description type. And then, the text structures of the types of r sequence, comparison, cause and effect, and problem solving were taught.

The teaching of the informative texts in the control group was carried out in accordance with the stages given in the teacher's guidebook of the Turkish lesson.

Data Analysis

Arithmetic mean, standard deviation, unpaired t test, and one-way analysis of variance were used in the analyses of the study data. The data obtained from the research was analysed through the use of SPSS 16.0 statistical package program. Raw scores acquired from the pre-tests and post-tests of the research were re-calculated over 100 points.

It was analysed with Kolmogorov-Smirnov test whether the scores acquired from the pre-tests and post-tests of the research indicated normal distribution or not. Since the data ranged between normal values, parametric tests were used.

At the stage of identifying the equivalences of the classes, one-way analysis of variance (ANOVA) was used to determine whether there was a significant difference between the scores gotten by different classes or not.

Unpaired t test was used to determine whether there was a significant difference between the pre-test and post-test scores of the students from the experimental group and the control group.

Results

Awareness of Informative Text Structures

It was tested through the unpaired *t*-test whether there was a significant difference between *"the informative text structure awareness"* of the students in the control group and of the students in the experimental group and the results are given in Table 2.

Table 2. *T-test Results of The Experimental Group and Control Group Students' Pre/Post Test Scores with Regard to The Informative Text Structure Awareness*

		n	*M*	SD	*t*	*p*
pre-test	EG	30	45.78	21.71	-1.126	0.265*
	CG	32	51.46	17.94		
post-test	EG	30	82.67	11.69	6.342	0.000**
	CG	32	53.33	22.66		

* $p>0.05$; ** $p<0.05$

As seen in Table 2, it is observed that the control group students' *level of knowledge with regard to the text structures* measured before the practice (M= 51.46) was higher compared to the experimental group students' levels (M= 45.78). However, there is not a significant difference between the experimental group and control group students' levels of text structure knowledge, which were measured before the practice ($p>0.05$).

It is observed that the experimental group students' *level of knowledge with regard to the text structures* measured after the practice (M= 82.67) was higher compared to the

control group students' level (*M*= 53.33). A significant difference on the behalf of the experimental group was found between the experimental group and control group students' levels of text structure knowledge, which were measured after the practice (p<0.05).

It was tested through the unpaired *t*-test whether there was a significant difference between the experimental group and control group students' reading comprehension levels, which were measured *before the practice*, and the results are given in Table 3.

Reading Comprehension Levels

Table 3. *T-test Results of The Experimental Group and Control Group Students' Pre-Test Scores with Regard to Their Levels of Reading Comprehension*

pre-test		*n*	*M*	SD	*t*	*p*
Description	EG	30	38.22	14.54	0.128	0.898*
	CG	32	37.71	16.82		
Sequence	EG	30	44.44	12.67	0.932	0.355
	CG	32	40.63	18.81		
Cause-effect	EG	30	49.11	18.98	0.252	0.802*
	CG	32	47.92	18.27		
Compare-contrast	EG	30	38.22	13.44	1.845	0.070*
	CG	32	31.88	13.62		
Problem solving	EG	30	33.78	15.46	-1.242	0.219*
	CG	32	38.54	14.74		
Reading Comprehension	EG	30	40.76	11.44	0.460	0.647*
	CG	32	39.33	12.81		

*p>0.05

When Table 3 is analysed, it is observed that the experimental group students' reading comprehension levels with regard to description (*M*= 38.22), sequence (*M*= 44.44), cause and effect (*M*= 49.11), and comparison (*M*= 38.22), which were measured before the practice, were higher compared to the levels of control group students. On the other hand, reading comprehension level of the control group students with regard to the problem solving (*M*= 38.54) was found higher compared to the level of experimental group students. However, there is not a statistically significant difference between the experimental group and control group students' levels of reading comprehension with regard to description, sequence, comparison, cause and effect, and problem solving skills, which were measured before the practice (p>0.05).

It was tested through the unpaired *t*-test whether there was a significant difference between the experimental group and control group students' reading comprehension levels which were measured *after the practice* and the results are given in Table 4.

Table 4. *T-test Results of The Experimental Group and Control Group Students' Post-Test Scores with Regard to Their Levels of Reading Comprehension*

post-test		*n*	*M*	SD	*t*	*p*
Description	EG	30	70.00	14.73	5.531	0.000*
	CG	32	46.25	18.70		
Sequence	EG	30	72.89	13.33	7.213	0.000*
	CG	32	43.13	18.55		
Cause-effect	EG	30	66.89	18.24	2.912	0.005*
	CG	32	53.33	18.39		
Compare-contrast	EG	30	63.55	16.40	4.244	0.000*
	CG	32	45.00	17.92		
Problem solving	EG	30	63.33	18.34	3.090	0.003*
	CG	32	47.50	21.73		
Reading Comprehension	EG	30	67.33	14.26	5.193	0.000*
	CG	32	47.04	16.36		

*p<0.05

When Table 4 is analysed, it is observed that the experimental group students' reading comprehension levels with regard to description (*M*= 70.00), sequence (*M*= 72.89), cause and effect (*M*= 66.89), comparison (*M*= 63.55), and problem solving (*M*= 63.33) which were measured after the practice, were higher compared to the levels of control group students. And statistically, a significant difference was found between the experimental group and control group students' levels of reading comprehension with regard to description,

sequence, comparison, cause and effect, and problem solving skills, which were measured after the practice (p<0.05).

General reading comprehension level of the experimental group students, measured after the practice, was found higher (M= 67.33) compared to the level of control group students (M= 47.04). And statistically, a significant difference was found between the experimental group and control group students' levels of general reading comprehension measured after the practice (p<0.05). The acquired difference is on the behalf of the experimental group.

When the post-test scores obtained by the experimental group students after the practice are analysed, it is observed that they succeeded most in the sequence type comprehension test (M= 72.89), and then in description (M= 70.00), cause and effect (M= 66.89), comparison (M= 63.55), and problem solving (M= 63.33) comprehension tests.

Discussion

According to the results of the research, reading comprehension skills of the students who learned the informative texts through the effective strategies developed in a positive way. These findings show similarity with the findings of the researchers who have taught informative text structures through various strategies (Baştuğ & Keskin, 2011; Dreher & Gray, 2009; Hall et al., 2005; Hoffman, 2010; Mcginley, 2008; Newman, 2007; Nubla-Kung, 2008; Özmen, 2011; Reutzel, Read & Fawson, 2009; Stagliano & Boon, 2009; Williams, 2005; Williams, 2008).

In the conclusions of the studies mentioned above, it is stated that the teaching of text structures has positive effects on reading comprehension. These findings have the quality to support the results of the research.

In the teaching of informative text through processual model, learning the structure of a text is the first stage. In the following stage, readers should choose important ideas given in the text and associate them with each other (Gunning, 2005). At this stage of the research, activities that would help students to make arrangements between the important ideas given in the text were used. With the help of these activities, students could organize supporting ideas and could more easily and systematically identify the main idea of the text, topic, and the aim of the writer.

According to the results of the research, the average value of the scores that students from the experimental group got from the *"sequence"* type, one of the reading comprehension tests applied after the practice, was higher compared to other tests. And, the types of *"description", "cause and effect",* and *"comparison"* followed it. The minimum point average was observed in reading comprehension test of the *"problem solving"* type. These results showed similarities with the findings that Sharp (2004) acquired from his study. In his study, Sharp (2004) assessed the informative text (description, sequence, cause and effect, and problem solving) comprehension skills of the students in the study group through gap-filling and recall tests. According to the results of the study, the students succeeded most in understanding the texts written in "sequence" type. And, texts written in *"description"* and *"cause and effect"* types followed it. The text type in which the reading comprehension level was the lowest was *"problem solving"* type.

In conclusion, the findings of this research show similarities with the findings of other researches in the literature. According to the results of the research, the teaching of informative texts through processual model develops the reading comprehension skills of students.

References

Akyol, H. (2007). Türkçe öğretim programı: Yenilikler ve sorunlar. *İlköğretmen Dergisi, 61*.

Akyol, H. (1999). Bilgi vermeye dayalı metinler ve öğretimi. *Çağdaş Eğitim Dergisi, 253,* 7-13.

Armbruster, B.B. (2004). Considerate texts. In D. Lapp, J. Flood and N. Farnan (Eds.), *Content area reading and learning: Instructional strategies (*pp. 47-58). Mahwah, NJ: Erlbaum.

Baştuğ, M. & Keskin, K. (2011). Bilgilendirici metin yapılarının öğretiminin okuduğunu anlamaya etkisi. *e-Journal of New World Sciences Academy, 6*(4), 2598-2610.

Cha, K. A. & Swaffar, J. (1998). The case for a procedural model as a multiple measure of reading comprehension. *System, 26*, 205-222.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.* London: Sage Publications.

Çakır, Ö. (l995). *Büyük ölçekli kuralların öğretiminin okuduğunu anlamaya etkisi,* (Unpublished doctoral dissertation). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Çıkrıkçı, S. (2004). *İlköğretim öğrencilerinde özetleme becerisinin gelişimi*, (Unpublished doctoral dissertation). Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Dönmez, C. & Yazıcı, K. (2006). Sosyal bilgilerde öğrencilerin okuduğunu anlama becerilerinin geliştirilmesinde metin yapısına bağlı olarak kullanılabilecek strateji ve teknikler. *Manas Üniversitesi Sosyal Bilimler Dergisi, 16*,138-154.

Dreher, M.J. & Gray, J.L. (2009). Compare, contrast, comprehend: using compare-contrast text structures with ells in K-3 classrooms. *Reading Teacher, 63*(2), 132-141.

Duke, N. K. & Pearson, P. (2002). Effective practices for developing reading comprehension. In A.E. Farstrup, & S. Samuels (Eds.), *What Research Has to Say About Reading Instruction* (pp. 205-242). Newark, DE: International Reading Association.

Dymock, S. (2005). Teaching expository text structure awareness. *The Reading Teacher, 59*(2), 177-181.

Görgen, İ. (1997). *Özetleme ve bilgi haritası oluşturma öğretiminin bilgilendirici bir metni öğrenme ve hatırlama düzeyine etkisi,* (Unpublished doctoral dissertation). Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

Gunning, T. G. (2005). *Creating literacy instruction for all students (5th ed.).* Boston: Allyn and Bacon.

Güneş, F. (2007). *Türkçe öğretimi ve zihinsel yapılandırma*. Ankara: Nobel Yayıncılık.

Hall, K.M., Sabey, B.L. & Mcclellan, M. (2005). Expository text comprehension: helping primary-grade teachers use expository texts to full advantage. *Reading Psychology, 26,* 211–234.

Hoffman, K. F. (2010). *The impact of graphic organizer and metacognitive monitoring instruction on expository science text comprehension in fifth grade students,* (Unpublished doctoral dissertation). University of North Carolina State, North Carolina.

Karatay, H. (2011). *Okuma eğitimi kuram ve uygulama*. Ankara: Berikan yayınevi.

Keçik, İ. (1993). *İlkokul 2. ve 5. sınıf öğrencilerinin hatırlama ve özet metinlerinde uygulanan büyük ölçekli yapı kuralları,* Dilbilim Araştırmaları, (p. 89-97), Ankara: Hitit Yayınevi.

Kintsch, W. & VanDijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85,* 363–394.

Kintsch, W. (2002). On the notions of theme and topic in psychological process models of text comprehension. İn M. Louwerse and W. van Peer (Eds.), *Thematics: Interdisciplinary Studies* (pp. 157-170). Amsterdam: Benjamins.

Kintsch W. & Kintsch, E. (2005). Comprehension. In S. G. Paris and S. A. Stahl (Eds.), *Current issues in reading comprehension and assessment* (pp. 71-92). Mahwah, NJ: Lawrence Erlbaum Associates.

Kuzu, T.S. (2003). *Etkileşimsel model'e uygun okuma öğretiminin Türkçe bilgilendirici metinleri anlama düzeyine etkisi*, (Unpublished doctoral dissertation). Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.

McGinley, D. L. (2008). *Effects of guided reading on motivation and comprehension in reading expository texts (*Unpublished graduate dissertation). Caldwell College, Program of Curriculum and Instruction.

Minskoff, (2005). *Teaching reading to struggling lerarners*. Maryland: Paul H. Brookes Publishing.

Moss, B. (2004). Teaching expository text structures through information trade book retellings. *The Reading Teacher, 57*(8), 710-718.

National Reading Panel. (2000). *Report of the National Reading Panel. Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Web: https://www.nichd.nih.gov/publications/pubs/nrp/Documents/report.pdf

Newman, L. M. (2007). *The effects of explicit instruction of expository text structure incorporating graphic organizers on the comprehension of third-grade students, (*Unpublished doctoral dissertation). University of Maryland.

Nubla-Kung, A. M. (2008). *The effects of text structure and signals on second graders' recall and comprehension of expository text, (*Unpublished doctoral dissertation). University of Columbia.

Özmen, R.G. (2011). Zihinsel yetersizliği olan öğrencilerin bilgi veren metinlerdeki bilgiler hatırlamalarında şematik düzenleyicilerin iki farklı sunum şeklinin karşılaştırılması. *Kuram ve uygulamada eğitim bilimleri, 11*(2), 785-793.

Read, S., Reutzel, D., & Fawson, P., (2008). Do you want to know what I learned? using informational trade books as models to teach text structure. *Early Childhood Education Journal, 36*(3), 213-219.

Rozmiarek, R. (2006). *İmproving reading skills across the content areas.* California: Corwin Press.

Schirmer, B. R. (2010). *Teaching the struggling reader.* Boston: Pearson-Allyn and Bacon.

Sharp, A. (2004). Strategies and predilections in reading expository text: The importance of text patterns. *RELC, 35* (3), 329-349.

Simonsen, S. (2004). Identifying and teaching text structures in content area classrooms. In D. Lapp, J. Flood, & N. Farnan (Eds.), *Content area reading and learning: Instructional strategies.* Mahwah, NJ: Erlbaum.

Stagliano, C. & Boon, R. T. (2009). The effects of a story-mapping procedure to improve the comprehension skills of expository text passages for elementary students with learning disabilities. *Learning Disabilities: A Contemporary Journal, 7*(2), 35-58.

Temple, C., Ogle, D., Crawford, A. & Freppon, P. (2005). *All children read: teaching for literacy in diverse classrooms*. New York: Addison Wesley Longman.

Tompkins, G.E. (2006). *Literacy for the 21st century: a balanced approach.* Upper Saddle River, NJ: Prentice Hall.

Weaver, C. A. and Kintsch, W. (1996). Expository Text. In R. Barr et al. (eds.), *Handbook of Reading Research* (Volume II, pp.230-245). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Williams, J.P. (2005). İnstruction in reading comprehension for primary-grade students: a focus on text structure. *The Journal Of Special Education, 39* (1), 6-18.

Williams, J. P. (2008). Explicit instruction can help primary students learn to comprehend expository text. İn C. C. Block and S. R. Parris (Eds.), *Comprehension processes: Research-based best practices,* (2nd edition). New York: Guilford Press.

Vacca, R. T. & Vacca, J. L. (2005). *Content area reading.* (Eighth Edition). Boston: Pearson Education.

This page is intentionally left blank

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

Structural Relationships Among Variables Affecting Elementary School Students' Career Preparation Behavior: Using a Multi-Group Structural Equation Approach

Sun Hee Park [a], JuSung Jun [b,*]

Received: 23 October 2017
Revised: 1 December 2017
Accepted: 5 December 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236122

Abstract

The purpose of this study was to analyze the structural relationships between parent support, career decision self-efficacy, career maturity, and career preparation behavior for elementary school students (5th and 6th grade) in Korea and to examine if there are gender differences. A total of 609 students of 7 elementary schools in Seoul, Korea was participated in this research. The collected data were analyzed using a structural equation model. The conclusions are as follows: First, parent support had significant positive effects on career preparation behavior but affected negatively career maturity. Second, career decision self-efficacy had a significant effect on career maturity and career preparation behavior for both boys and girls in elementary school. Finally, career decision self-efficacy had a more direct and indirect influence on career maturity and career preparation behavior than parent support.

Keywords: Career preparation behavior, parent support, career decision self-efficacy, career maturity.

Introduction

As the Fourth Industrial Revolution unfolds, transforming and disrupting many existing models and processes, the job market has come to be characterized by instability and unpredictability. Resultantly career education, especially for adolescents, is now a subject of great importance. Some developed countries are developing and operating career development programs from the elementary school level in order to prepare for such a future society (Coogan, 2016). Recognizing this, the Korean government enacted the Career Education Act (June 22, 2015) and the Enforcement Decree of the Career Education Act (December 22, 2015) to establish the basic direction and a support system for career education, which included putting in place national career education goals and achievement standards. The government has also placed specialist career teachers in secondary schools and introduced a system for career education in primary and secondary schools.

The 2015 amendment specifies the goals of elementary school career education as being creating a positive self-concept, recognizing the importance of employment, and fostering basic career development skills through developing key skill-sets such as job searching, planning and preparation (Ministry of Education, 2015). Career education in elementary schools focuses on the career recognition stage. In particular, the 5th and 6th grades of elementary school are seen as a vital time for students to consider and even select their own careers, and a time when it is necessary to provide opportunities for systematic career searching (Lee, Lee, & Lee, 2016). In recent years, several studies (Ahn, 2008; Chung, 2014; Gushue & Whitson, 2006; Han & Cho, 2016; Han & Oh, 2014; Seo, 2016) have been conducted on the variables that affect career preparation behavior and how they fit into career education for adolescents. In particular, variables that have been studied extensively in career preparation behavior research include parent support, career decision self-efficacy, and career maturity. Career preparation behavior is usually understood as the concrete and practical behavioral efforts undertaken in the process of achieving career goals. Career preparation behavior should be undertaken as part of the process of finding a reasonable and suitable career, and it includes all actions to implement a decision about a career after that decision is reached (Kim, 1997). Parent support is an environmental variable that represents the perception of children about the way that their parents provide educational and occupational efficacy information (Ahn, 2008). Parent support is information that allows children to believe that their parents are caring and loving, that they are proud of their children, and that they are members of an accountable communication network (Cobb, 1976). Following Ahn (2008), in this study, the researchers define career-related parent support as the perception of children about how their parents provide educational and occupational efficacy information. Hackett and Betz (1981) applied Bandura's concept of self-efficacy to career development theory as an important cognitive variable affecting career decision process. They conceptualized 'career self-efficacy' specifically for the individual's efficacy related to career choice and adaptation. Career decision self-efficacy can be defined as beliefs about the individual's ability to successfully perform tasks related to career decisions

[a] Seoul Semyung Elementary School, E-mail: chrysta@sen.go.kr
[b,*] Corresponding author: JuSung Jun, Soongsil University, Department of Lifelong Education, 369 Sangdo-Ro, Dongjak-Gu, Seoul, 06978, Korea. E-mail: jnet@ssu.ac.kr

according to the definition of Hackett and Betz (1981). Career maturity is defined as the degree to which an individual understands and evaluates himself/herself and selects and adapts his/her career path (Chung, 2014).

Previous studies demonstrate positive relationships between the variables of parent support, career decision self-efficacy, career maturity, and career preparation behavior. Firstly, Gushue and Whitson (2006) and Pečiulytė, Ustinavičiūtė, and Norvilė (2014) found that there was a positive relationship between high school student's parent support and career decision self-efficacy (Pečiulytė, Ustinavičiūtė, & Norvilė, 2014). Secondly, Han and Oh (2014) indicated that parent support had a significant effect on the levels of career maturity of high school students. Choi (2016) also found that middle school students who perceived parent support levels as high were highly confident about their careers, and were able to independently undertake career preparation and determine their career paths. In other words, parent support had a positive impact on students' career maturity. Thirdly, Ahn (2008) and Han and Cho (2016) showed that parent support has a positive effect on career preparation behavior. Fourthly, Lee and Cho (2015) and Seo (2016) showed that college students' career maturity had a significant effect on their career preparation behavior. Fifthly, Walker (2010), Lee and Lee (2000) found that career decision self-efficacy of college students had a positive effect on career maturity. Finally, Kim and Lee (2013) reported that the higher the levels of career decision self-efficacy of middle and high school students, the more career preparation behavior they engaged in. Seo (2016) also indicated that career decision self-efficacy had a positive effect on the career preparation behavior of college students.

However, the above studies all focused only on high school and college students, and also sought to identify only simple relationships between two of the variables in question. In other words, existing studies have had a major interest in identifying variables that affect career preparation behavior of high school and college students. This study attempts to analyze the multiple and causal relationships among the main variables that affect career preparation behavior through a single structural equation model. In the context of the importance of career education for elementary students, it is necessary to analyze the structural influences of the variables.

The purpose of this study is to analyze the structural relationships between the variables of parent support, career decision self-efficacy, career maturity, and career preparation behavior for elementary school students (5th and 6th grade), and further to examine if there are gender differences in these relationships. The results of this study can contribute not only to theorizing the multifaceted relationships among the variables that influence the level of career preparation behavior of elementary students, but also to suggesting the implications for elementary students' career research.

Methodology

Research Model

In this study, the researchers derived a research model(see Figure 1) based on the results of previous studies on parent support, career decision self-efficacy, career maturity, and career preparation behavior.

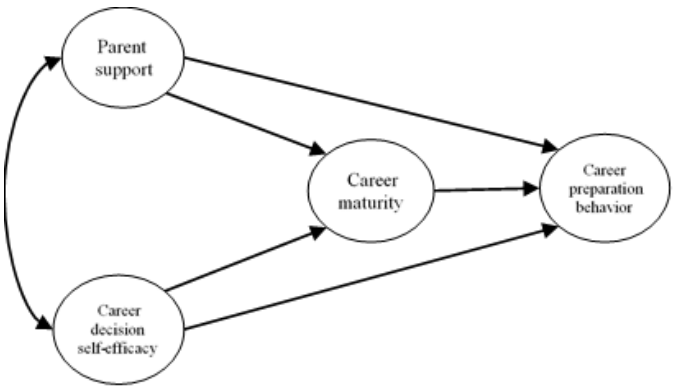


Figure 1. Research model

Research based on this type of design aids researchers and educators to better understand how to orchestrate learning experiences among children in a daily educational context, and at the same time, how to develop theoretical ideas regarding the nature of learning (Bell, 2004). Consequently, the research design takes into consideration primarily understanding of the processes and the strategies that help learners to be successful in solving math word problems.

Research Samples and Data Collection

This study was conducted on 5th and 6th grade students at seven elementary schools in Seoul, Korea. A total of 638 questionnaires were collected and 609 were used (responses deemed untruthful were excluded). Listwise deletion was used for handling missing data. Of the respondents, 304 (49.9%) were boys and 305 (50.1%) were girls. There were 324 students (53.2%) in the 6th grade and 285 students (46.8%) in the 5th grade.

Measurement Scales

Parent support scale. For the parent support scale the 'Career-related Parent Support Scale' developed by Tuner, Annette, Lapan, Udipi, and Ergun (2003) and translated into Korean by Kim (2004) was used. The scale consists of the four sub-factors of career-related modeling,

emotional support, instrumental support, and verbal support. It has a total of 27 items and uses a 5-point-Likert scale format. In the present study, the whole Cronbach α for the parent support scale was .94, and Cronbach α for the sub-factors was career-related modeling .89, emotional support .88, instrumental support .85, and verbal support .81.

Career decision self-efficacy scale. For this scale the 'Career Decision-Making Self-Efficacy Scale, CDMSES' developed by Taylor and Betz (1983) and validated by Lee (2001) for middle and high school students was used. This scale consists of the four sub-factors of goal selection, job information, problem solving, and future planning. It has a total of 25 items and uses a 5-point Likert scale format. In this study, the overall Cronbach α for the career decision self-efficacy scale was .95, and Cronbach α for the sub-factors was goal selection .92, job information .85, problem solving .73, and future planning .83.

Career maturity. In this study, the researchers used the career maturity measurement tool developed by the National Youth Policy Institute (2010). It has a total of 7 items and uses a 5-point-Likert scale format. The Cronbach α of the seven items was .72.

Career preparation behavior scale. The career preparation behavior scale was developed by Choi, Kim, Hwang, and Huh (2009) in order to measure the career preparation behavior of youth. This scale consists of the two sub-factors of career exploration and career development, and includes 10 items. A Likert type 5 point scale was used. In this study, the overall Cronbach α for the career preparation behavior scale was .84, and Cronbach α for the sub-factors was career exploration .74 and career development .76 respectively.

Data Analysis

In order to investigate the normal distribution of observational variables, descriptive statistics such as mean, standard deviation, skewness, and kurtosis of the variables were analyzed using SPSS. In addition, the fitness of the research model was estimated by applying the maximum likelihood method (ML) using AMOS.

Results

Descriptive Statistics and Correlations

Table 1 revealed that the descriptive statistics of the measurement variables are not problematic for the normality assumption. Table 2 represents the correlation matric between variables.

Table 1. *Descriptive Statistics (N=609)*

Latent variables	Observed variables	Total(609)		Boys(304)		Girls(305)		Skewness	Kurtosis
		M	SD	M	SD	M	SD		
Career decision self-efficacy	Goal selection	3.780	.772	3.774	.809	3.786	.734	-.502	.253
	Job information	3.492	.838	3.522	.880	3.462	.795	-.158	-.135
	Problem solving	3.463	.861	3.467	.932	3.459	.785	-.074	-.274
	Future planning	3.441	.852	3.497	.895	3.385	.806	-.114	-.148
Parent support	Modeling	3.754	.985	3.751	1.013	3.756	.958	-.620	-.293
	Emotional support	3.451	.924	3.424	.945	3.477	.903	-.417	-.155
	Instrumental support	3.140	.887	3.156	.897	3.123	.877	-.242	-.322
	Verbal support	3.671	.909	3.722	.914	3.620	.903	-.540	-.136
Career preparation behavior	Career exploration	3.057	.875	3.015	.911	3.098	.838	.030	-.434
	Career development	3.619	.940	3.566	.984	3.672	.893	-.494	-.269
Career maturity	Career maturity	3.812	.710	3.803	.715	3.822	.707	-.253	-.322

Table 2. *Correlations between Variables*

	Career decision self-efficacy				Parent support				Career preparation behavior		Career maturity
	1	2	3	4	5	6	7	8	9	10	11
1	1										
2	.792**	1									
3	.460**	.500**	1								
4	.838**	.802**	.483**	1							
5	.406**	.433**	.314**	.390**	1						

Table 2 (Cont.). *Correlations between Variables*

6	.543**	.540**	.319**	.549**	.471**	1					
7	.518**	.532**	.326**	.556**	.498**	.750**	1				
8	.497**	.486**	.363**	.485**	.430**	.724**	.633**	1			
9	.584**	.575**	.307**	.536**	.401**	.445**	.470**	.326**	1		
10	.659**	.581**	.291**	.608**	.340**	.459**	.497**	.357**	.655**	1	
11	.508**	.404**	.083*	.475**	.199**	.215**	.286**	.119**	.341**	.409**	1

** p<.01

Note: ① Goal selection ② Job information ③ Problem solving ④ Future planning ⑤ Modeling ⑥ Emotional support ⑦ Instrumental support ⑧ Verbal support ⑨ Career exploration ⑩ Career development ⑪ Career maturity

Structural Relationships of the Research Model

Verification of model fit. The model fit of the initial model is χ^2=196.804 (df=39, p<.001), TLI (Tucker-Lewis Index)=.947, CFI (Comparative Fit Index) =.962, and RMSEA(Root Mean Square Error of Approximation) =.082. The values of TLI and CFI are over .90, which indicates that the initial structural model meets the criteria for fitness. Therefore, the initial model was selected as the final model without modification.

The results of the standardized model estimation of the final structural regression model are shown in Figure 2, and the statistical significance test results are shown in Table 3. Additionally, the indirect, direct, and total effects of path coefficients between the variables are shown in Table 4.

Looking at the path coefficients between variables, parent support (β=-.19) had a negative impact on career maturity, and career decision self-efficacy (β=.64) had a significant positive effect on career maturity. In addition, parent support (β=.17) and career decision self-efficacy (β=.64) had a significant effect on career preparation behavior. The total effect of parent support and career decision self-efficacy on career preparation behavior was .152, .701, respectively. Career decision self-efficacy was found to be more influential on career preparation behavior than parent support.

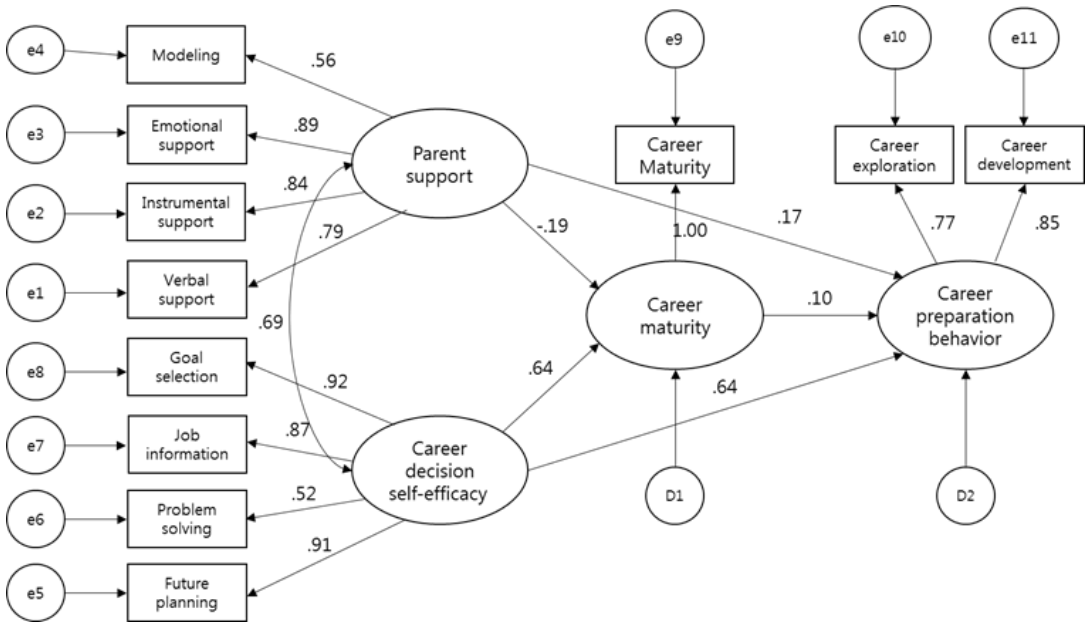


Figure 2. The structural regression model (standardized coefficient)

Table 3. *Parameter Estimation and Significance Verification of the Structural Regression Model*

Path	Unstandardized estimate	Standardized estimate	S.E.	C.R
Parent support → Career maturity	-.186	-.188	.056	-3.331***
Career decision self-efficacy → Career maturity	.585	.641	.051	11.436***
Career maturity → Career preparation behavior	.094	.098	.037	2.562**
Career decision self-efficacy → Career preparation behavior	.555	.638	.054	10.315***
Parent support → Career preparation behavior	.161	.171	.049	3.283**

** p<.01, *** p<.001

Table 4. *Direct Effect, Indirect Effect, Total Effect of Path Coefficients between Latent Variables*

Path	Direct effect	Indirect effect	Total effect
Parent support → Career maturity	-.188	-	-.188
Career decision self-efficacy → Career maturity	.641	-	.641
Career maturity → Career preparation behavior	.098	-	.098
Parent support → Career preparation behavior	.171	-.018	.152
Career decision self-efficacy → Career preparation behavior	.638	.063	.701

Table 5. *Fit Indices for Measurement Invariance Test*

Model	χ^2	df	CFI	TLI	RMSEA	$\Delta\chi^2$	Sig.Dif
[Model 1] Unconstrained	214.25	78	.774	.682	.054		
[Model 2] λ constrained	216.31	85	.782	.718	.050	2.06	No
[Model 3] Φ constrained	264.91	88	.958	.948	.058	50.66	Yes
[Model 4] λ,Φ constrained	231.05	93	.771	.729	.049	16.8	No
[Model 5] λ,Φ,θ constrained	238.09	105	.779	.769	.046	23.84	No

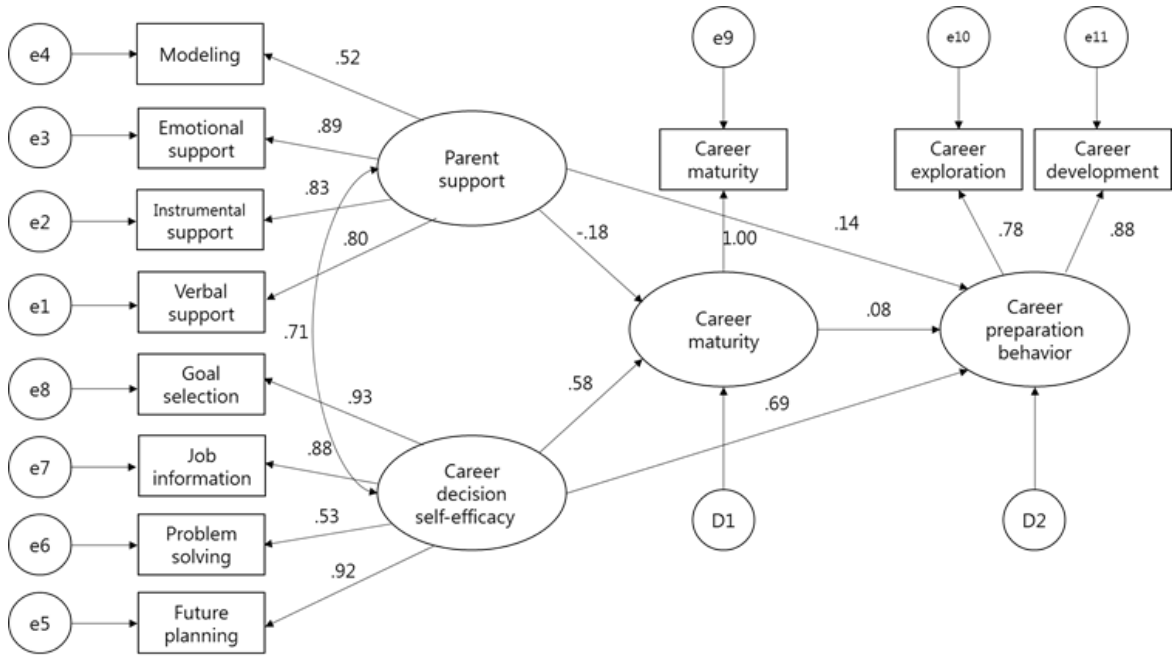
Multiple group structural regression models for gender

Measurement invariance verification. In order to conduct the multiple group path analysis for boys and girls, measurement invariance should be assured that the factor loadings of these groups are the same. The result of the analysis shows that there were no problems with measurement invariance (See Table 5).

Multiple group path analysis for gender

Figure 3 indicates the results of a multiple group path analysis for gender. Table 6 represents path coefficients comparison for gender between the variables. The indirect effect, the direct effect, and the total effect of the path coefficients for gender between the variables are shown in Table 7. Both boy and girl students had higher career decision self-efficacy influence on career preparation behavior than parent support had on career preparation behavior. The effect of parent support on career maturity was negatively greater for girls (β=-.207) than for boys (β=-.183). The effect of career decision self-efficacy on career maturity was higher in girls (β=.726) than boys (β=.577). The effect of parent support on career preparation behavior was higher for girls (β=.20) than for boys (β=.14). In addition, the effect of career decision self-efficacy on career preparation behavior was higher in boys (β=.689) than girls (β=.578).

Boys



Girls

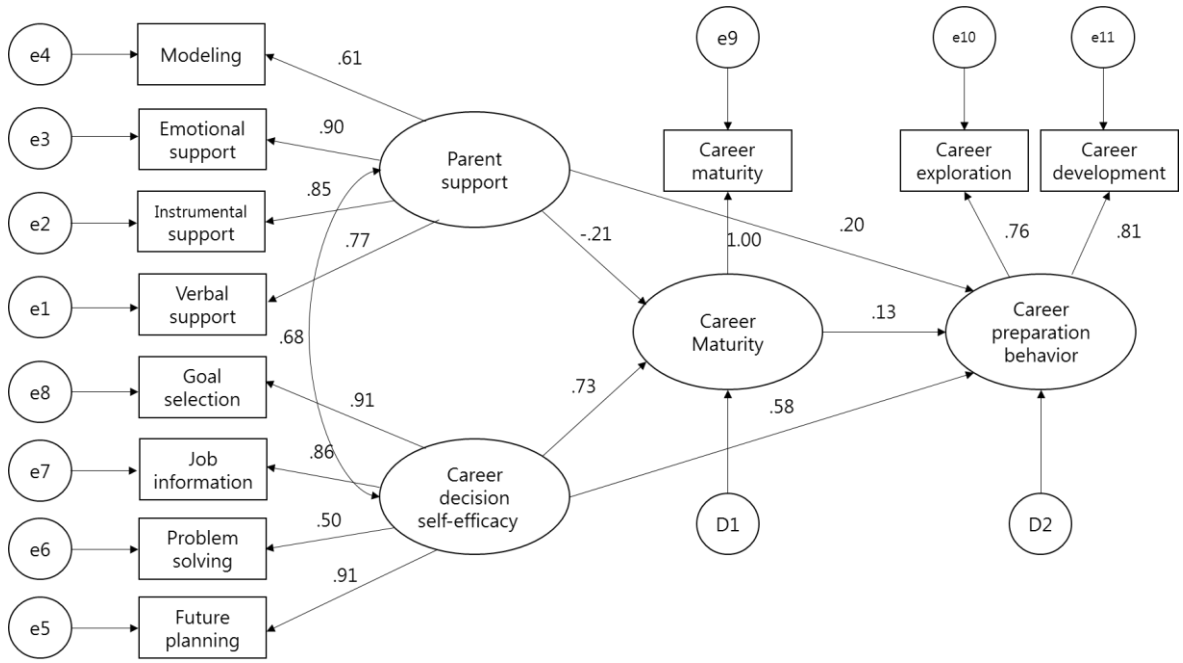


***Figure* 3.** Multiple group structural regression models for gender

Table 6. *Path Coefficients Comparison for Gender*

Path	Boys		Girls	
	Standardized estimates	C.R.	Standardized estimates	C.R.
Parent support → Career maturity	-.183	-2.191*	-.207	-2.766**
Career decision self-efficacy → Career maturity	.577	6.988***	.726	9.657***
Career maturity → Career preparation behavior	.076	1.542	.132	2.136*
Career decision self-efficacy → Career preparation behavior	.689	8.328***	.578	6.010***
Parent support → Career preparation behavior	.142	2.025*	.201	2.605**

*p<.05, **p<.01, ***p<.001

Table 7. *Indirect, Direct, and Total Effects of the Path Coefficients for Gender between Variables*

Path	Direct effect		Indirect effect		Total effect	
	Boys	Girls	Boys	Girls	Boys	Girls
Parent support → Career maturity	-183	-.207	-	-	-.183	-.207
Career decision self-efficacy → Career maturity	.577	.726	-	-	.577	.726
Career maturity → Career preparation behavior	.076	.132	-	-	.076	.132
Parent support → Career preparation behavior	.142	.201	-.014	-.027	.129	.173
Career decision self-efficacy → Career preparation behavior	.689	.578	.044	.096	.733	.674

The moderating effect of gender is in <Table 8>. The effect of career decision self-efficacy on career maturity was only significant: The path coefficient was higher in girls (β=.726) than boys (β=.577).

Table 8. *The Comparison for $\Delta\chi^2$ on the Path Coefficients*

Path	χ^2	df	χ^2/df	$\Delta\chi^2$	Sig.Dif
Unconstrained Model	248.996	78	3.192		
Parent support → Career maturity	249.067	79	3.153	.07	No
Career decision self-efficacy → Career maturity	252.845	79	3.202	3.85	Yes
Career maturity → Career preparation behavior	249.338	79	3.156	.34	No
Career decision self-efficacy → Career preparation behavior	249.719	79	3.161	.72	No
Parent support → Career preparation behavior	249.189	79	3.154	.19	No

Discussion and Conclusion

The results of this study showed that parent support perceived by elementary school seniors had a negative effect on career maturity and a positive effect on career preparation behavior, respectively. In addition, parent support indirectly had a significant positive impact on career preparation behavior via career maturity. These results are consistent with the findings (Kim, 2004) that parent career support for middle and high school students had a positive correlation with career preparation behavior but did not directly lead to career maturity. On the other hand, these findings differ from those of Han and Oh (2014) and Choi (2016). These studies have shown that parent support positively influenced career maturity in middle and high school students. This study reveals that parent support may help students in the upper grades of elementary school to navigate and develop the career path, but it may interfere with career maturity if parents are involved excessively. This suggests that in order to strengthen the voluntary career preparation behavior of the elementary school students, the parents need to support their children's opinions and actions with appropriate encouragement while forming cooperative relationships with their children.

The career decision self-efficacy of the upper grades of elementary school showed a direct influence on career maturity and career preparation behavior. In addition, career decision self-efficacy indirectly had a significant positive influence on career preparation behavior via career maturity. These results are consistent with the findings (Cho, 2007; Kim, 2005) that the higher the career decision self-efficacy of college students, the higher the level of career preparation behavior. This study implies that career decision self-efficacy improvement programs encouraging elementary school students to set career goals, collect job information, and solve career-related problems themselves can promote their career preparation behavior. In addition, the direct and indirect effects of career decision self-efficacy on career preparation behavior were more significant than the effects of career-related parent support. This suggests that providing students who do not adequately receive career-related parent support with programs that enhance career decision self-efficacy can promote their career preparation behavior.

In comparison between gender groups, both boy and girl students had higher effects of career decision self-efficacy on career preparation behavior rather than the effects of parent support. In the moderating effect of gender, only the path between career decision self-efficacy and career maturity was statistically significant ($p<.05$). That is, girl students were higher than boy students in terms of the path coefficient. It is difficult to find studies that analyzed the differences between male and female students concerning the relationships between parent support, career decision self-efficacy, career maturity, and career preparation behavior. This study is significant in that it analyzed the difference in gender on the relationship between these variables. This shows that the results of the study may provide a basis for taking different strategies between boys and girls in providing educational implications.

The conclusions of this study are as follows:

First, parent support had significant positive effects on career preparation behavior but affected negatively career maturity. The results indicate that excessive career-related parenting may have a negative effect on the career maturity of the student. Therefore, it is important that parents form a cooperative relationship with their child from the time the child is in elementary school, and encourage and support their child to voluntarily and independently make decisions. In terms of educational policy, it is also necessary to provide education for parents so that they are aware of how to have a positive role in their child's career guidance. Second, career decision self-efficacy had a significant effect on career maturity and career preparation behavior for both boys and girls in elementary school. This implies that the development and application of career decision self-efficacy improvement programs in elementary education can have a positive effect on the career maturity and career preparation behavior of elementary students. In particular, gender difference analysis showed that the effect of career decision self-efficacy on career maturity was higher in girl students than in boys. This means that career decision self-efficacy improvement programs can be more positive for female students. Finally, career decision self-efficacy had a more direct and indirect influence on career maturity and career preparation behavior than parent support. This implies that active educational intervention is possible to improve the career decision self-efficacy of students.

Career education that equips students with the tools to help them prepare for and adjust to the unpredictable society that they will be adults in is now more important than ever. Up to now, research on career preparation behavior, a key variable in career development, has been mainly conducted for high school students and college students. This study is meaningful as it is a comprehensive analysis of the structural relations among the variables related to career preparation behavior for the upper grades of elementary school students. It is expected that systematic career education will be realized by developing and applying various career decision self-efficacy improvement programs to strengthen career preparation behavior from elementary school onwards. Parent education programs improving parent support are also needed and should be developed and implemented so that students' career preparation behavior can be improved from elementary school.

References

Ahn, H. K. (2008). *The influences of hope, parent support and career barrier on high school student's career preparation behavior* (Master's thesis, Dongguk University, Seoul, Korea).

Cho, M. S. (2007). The influence of ego-identity, social support, career decision-making self-efficacy on career maturity and career preparation behavior of

late adolescence. (Doctoral thesis, Sungkyunkwan University, Seoul, Korea).

Choi, I. J., Kim, B. H., Hwang, M. H., & Huh, E. Y. (2009). *Korea youth indicators survey 4* (Youth career occupation index). Korea Youth Policy Institute.

Choi, M. K. (2016). The relationships between negative affect, social support, and career maturity among adolescents. *Korean Journal of Child Studies*, *37*(3), 53-68. doi:10.5723/kjcs. 2016.37.3.53

Chung, M. J. (2014). The effects of the satisfaction in major and career maturity on the employment anxiety and employment preparation behavior of undergraduates majoring in airline service. *Journal of Korean Airline Management, 12*(1), 81-104.

Cobb, S. (1976). Social support as a moderator of life stress. *Psychosom Med, 38*, 300-314.

Coogan, T. A. (2016). Supporting school counseling in Belize: Establishing a middle school career development program. *International Electronic Journal of Elementary, 8*(3), 379-390.

Gushue, G. V., & Whitson, M. L. (2006). The relationship among support, ethnic identity, career decision self-efficacy, and outcome expectations in African American high school students: Applying social cognitive career theory. *Journal of Career Development, 33*(2), 112-124. doi:10.1177/ 089484530 6293416

Hackett, G., & Betz, N, E. (1981). A self-efficacy approach to the career development of women. *Journal of vocational behavior, 18*(3), 326-339.

Han, C. H., & Cho, Y. A. (2016). The relationship between parent career support and career preparation behavior among high school students: Mediating effects of self-encouragement. *Korean Journal of Youth Studies*, *23*(4), 79-107. doi:10.21509/KJYS.2016. 04.23.4.79

Han, S. M., & Oh, I. S. (2014). The influence of high school students` perceived parent support and career self-efficacy upon career maturity. *Secondary Education Research*, *62*(1), 261-281

Han, S. M., & Oh, I. S. (2014). The influence of high school students` perceived parent support and career self-efficacy upon career maturity. *Secondary Education Research*, *62*(1), 261-281

Jung, Y. H. (2016). Exploration of career education competency of elementary school teachers. *Global Creative Leader*, *6*(2), 49-67.

Kim, S. J. (2005). The influence of career barriers, hope and career decision-making self-efficacy on the career preparation behavior of college students. (Doctoral thesis, Hongik University, Seoul, Korea).

Kim, S. J., & Lee, J. J. (2013). The effect of career barriers and career decision-making self-efficacy on the career preparation behavior: For middle and high school and specialized high school students. *Research on Rehabilitation Psychology, 20*(1), 119-136

Kim, S. R. (2004). The influence of parent support, career decision-making self-efficacy and dysfunctional career thoughts on adolescents' career development. (Doctoral thesis, Hongik University, Seoul, Korea).

Lee, E. K. (2001). The effect of self-efficacy on career development. (Doctoral thesis, Ewha Woman University, Seoul, Korea).

Lee, H. R., Lee, N. H., & Lee, Y. O. (2016). The effects of career group counseling program on elementary school students' career consciousness maturity and human relations. *Child Education, 15*(1), 221-233.

Lee, G. H., & Lee, H. J. (2000). The effects of career self-efficacy in predicting the level of career attitude maturity of college students. *Korean Psychological Association: Counseling and Psychotherapy, 12*(1), 127-136.

Lee, J. M., & Cho, B. R. (2015). Structural relationships among creative personality, university supporting, career maturity, career preparation behavior of university students. *Creativity Education Research, 15*(4), 41-59

Ministry of Education (2015). Goals of school career education and achievement criteria (commentary). Korea Institute of Vocational Training.

Pečiulytė, E., Ustinavičiūtė, L., & Norvilė, N. (2014). The relationship between parental involvement and students' career decision self-efficacy. *Social Work Research Journal, 13*(1), 110-122. doi:10.13165/SD-14-13-1-10

Seo, K. H. (2016). The effects of career decision-making self-efficacy and career maturity on career preparation behavior of culinary and foodservice management major students. *Culinary Science & Hospitality Research*, *22*(3), 121-138. doi:10.20878/ cshr.2016.22.3.011

Taylor, K. M., & Betz, N. E. (1983). Application of self-efficacy theory to the understanding and treatment of career indecision. *Journal of Vocational Behavior, 22*(1), 63-81.

Tuner, S. S., Annette, A. B., Lapan, R. T., Udipi, S., & Ergun, D. (2003). The Career-related parent support scale. *Measurement and Evaluation in Counseling and Development*, *36*(2), 83-94.

Walker, Q. D. (2010). An investigation of the relationship between career maturity, career decision self-efficacy, and self-advocacy of college students with and without disabilities (Doctoral dissertation, University of Iowa, Iowa City, Iowa). Retrieved from: http://ir.uiowa.edu/cgi/ viewcontent. cgi?article=1802 &context=etd

iejee | INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effect of Primary School Students' Writing Attitudes and Writing Self-Efficacy Beliefs on Their Summary Writing Achievement

Pınar Bulut [a]

Received: 14 August 2017
Revised: 06 Sept. 2017
Accepted: 18 Sept. 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236123

Abstract

In this study, the effect of writing attitude and writing self-efficacy beliefs on the summarization achievement of the 4th grade primary school students was examined using the structural equation modeling. The study employed the relational survey model. The study group constructed by means of simple random sampling method is comprised of 335 fourth grade primary school students. In the collection of the research data, "Writing Self-efficacy Scale (Güneş, Kuşdemir and Bulut, 2017)", "Writing Attitude Scale (Can, 2016)" and "Summary Evaluation Form for Narrative Texts" (Bulut, 2013) were used. For the current study, the constructed hypothesis is "writing attitude and writing self-efficacy positively and significantly affect writing achievement" and thus a model was formed. In the analysis of the data, SPSS 22 and AMOS 22.0 program packages were used. As a result of the study, it was found that writing attitude and writing self-efficacy beliefs are highly correlated to each other and writing attitude and writing self-efficacy beliefs directly and significantly affect summary writing.

Keywords: Attitude, self-efficacy, writing attitude, writing self-efficacy, summary writing.

Introduction

Writing is an important expressive language skill that the individual performs using symbols that refer to voices. Writing is the most recently acquired skill that begins to be learned by reading during the formal education process. According to Güneş (2007), writing starts with the review of the information that is structured in the brain. Writing comes into being as a result of organization of the information to be selected by considering writing purpose, method, subject and limitations through various mental processes such as sorting, classifying, establishing relations, criticizing, estimating, analyzing and synthesizing and then expression of this organized information with letters, syllables, words and sentences.

Graham (2006; cited by Graham, Hebert, Sandbank & Harris, 2014) states that writing is a preferred tool for students to express what they know in today's classrooms. On the other hand, according to Troia and Graham (2003), writing is one of the most complicated literacy activities for adults and children.

Though writing is viewed to be a difficult language skill, there are many benefits incurred by the acquisition of this skill. Writing is a skill necessary for critical thinking, learning and expression. Students who develop effective writing skills during school years have a significant advantage over those who do not in their later life (Graham & Perin, 2007). In addition, because writing is interwoven with mental processes, it helps students to expand their thinking, organize their knowledge, use language, enrich their knowledge, and develop their mental dictionaries (Güneş, 2007). In addition to these benefits, the development of writing skills promotes the aesthetic sensitivity of students and then written expression starts to be perceived as an artistic endeavor and feeling rather than a science (Ungan, 2007).

There are many factors affecting the individual's writing achievement. They can be cognitive, affective or physical factors. The most important affective factors affecting writing achievement are attitude and self-efficacy belief.

Attitude involves an assessment of an object, person or event on a continuum extending from negative to positive and makes us prone to behaving in a certain way in the face of that object, person or event (Plotnik, 2009).

Göçer (2014) points out that the interests, desires and attitudes of each individual are not the same for the writing skills which are very important in the communication and social life of individuals and stresses that one of the factors that affect the written expression achievement of individuals and cause them to move away from writing is their attitude towards writing. Variables such as education, family environment, and personal characteristics are influential in the emergence of different attitudes towards writing (Akaydın and Kurnaz, 2015).

Graham, Berninger, and Fan (2007) define the writing attitude as an effective regulation that causes the author to feel happy or unhappy during the writing activity. According to Graham et al. (2007), students' writing attitudes affect their writing achievement.

Another affective factor affecting writing achievement is self-efficacy belief. Self-efficacy beliefs are defined as "people's judgments about their skills to organize and

[a] Correspondence details: Gazi University, Faculty of Education, Department of Elementary Education, Yenimahalle, Ankara/Turkey. E-mail: pnrbulut06@gmail.com (P. Bulut)

perform acts that are required to achieve certain types of performance" (Bandura, 1986; cited by Schunk, 1989). Senemoğlu (2005) defines the concept of self-efficacy as the individual's own judgment, belief about himself/herself regarding the extent to which he/she will be successful in overcoming difficult situations to be confronted with in the future. Self-efficacy is a key mechanism in social cognitive theory that assumes that achievements are based on interactions between behaviors, personal factors, and environmental conditions (Schunk, 2003). Self-efficacy is a key concept for ensuring student participation and encouraging learning (Graham, Harris, Bartlett, Popadopoulou & Santoro, 2014). Jinks and Lorsbach (2003) argue that self-efficacy may be a powerful tool for educators to meet the learning needs of students. According to Schunk (2003), at the beginning of learning activities, there is a sense of self-efficacy for students to achieve their goals. The student's self-evaluation of his/her learning process helps to protect self-efficacy and motivation.

The self-efficacy beliefs affect students' performance by influencing the choices made, the effort spent and the perseverance and anxiety felt in the face of distress (Pajares& Valiante, 1997). If the individual has low self-efficacy belief, he/she may feel insufficient in the relevant work or activity. Therefore, students' self-efficacy beliefs need to be nurtured. Senemoğlu (2005) makes some suggestions to the teachers in order to strengthen their students' self-efficacy perceptions. These are teaching in such a way as to meet the individual needs of students, conducting various activities so that each student can get engaged in, using cooperation-based teaching approaches and avoiding approaches based on the comparison of students. In Akar (2008), some suggestions are made to improve self-efficacy. Some of them are general and some others are directly related to writing. These are, feedback, setting goal, modeling, strategy use, assigning challenging tasks, positive high expectations, teaching methods, successful learning experiences, right to choice and teacher attitude.

There are different strategies and techniques that readers can apply at different stages of the reading process in order to make the reading process more efficient and to obtain the highest level of meaning from the reading. One of them is writing summary. Summarizing refers to the reader's finding the main idea in the text, removing unnecessary details and shortening the text without distorting the structure of the text and flow of thought with his/her own words and sentences by using some strategies (Bulut, 2013). Summarization is a multifaceted process that requires a variety of prerequisite skills of which finding the main idea is the most important and that enhances students' comprehension and cognitive abilities (Williams, 2007).

Summary writing is not a random writing exercise. Therefore, it should be taught to students by teachers. According to Anderson, Suzanne and Babadoğan (1991), in this teaching process, teachers must first choose a short and easy text. Later, longer texts can be used as students gain experience in summarizing. Summary writing must first begin with texts in the narrative type.

A qualified summary must have certain characteristics in terms of content, form, language and narrative. In addition, Brown, Day and Jones (1983) emphasize that there are five basic rules for producing a good summary. These are removal of insignificant information, removal of unnecessarily repeated information, generalization, selection of the topic sentence and originality. These rules can be regarded as summarizing strategies.

There is research showing that there is a relationship between students' attitudes towards writing, writing self-efficacy and writing achievement. Graham, Berninger and Fan (2007), in their study investigating the relationship between the primary school students' writing achievement and writing attitude, concluded that age factor is important and the third-grade students are more successful in writing than the first-grade students. In the same study, the girls were found to have more positive attitudes towards writing than the boys. However, gender was not found to be significantly influential on writing achievement. The researchers stated that differences in the students' writing attitudes and behaviors led to individual differences in writing achievement. They found that the students having more positive attitudes towards writing had higher writing achievement than their counterparts with weaker attitudes towards writing. Demir (2013) found a positive, medium and statistically significant correlation between the secondary school 8th grade students' creative writing competence and writing self-efficacy perceptions. It was stated that in general students with higher self-efficacy have better creative writing skills. Baştuğ (2015), in a study focusing on writing and affective factors affecting it, found that writing disposition and attitude significantly and positively affect writing achievement; writing attitude and disposition significantly and negatively affect writer's block; writer's block negatively and significantly affects writing achievement. It was found that writer's block has the greatest effect on writing achievement, followed by writing disposition and writing attitude.

The level of schooling was also found to be effective on writing attitude. Yıldız and Kaman (2016), found that while the primary school students' attitude towards writing is significantly higher than that of the secondary school students, their attitudes do not vary significantly depending on gender.

Writing achievement depends on writing self-efficacy and writing attitude. For achievement, self-efficacy level should be high. Moreover, the attitude developed towards writing directly affects achievement. The current study investigated the direct effect of writing attitude and writing self-efficacy on writing achievement. To this end, the following model was constructed.

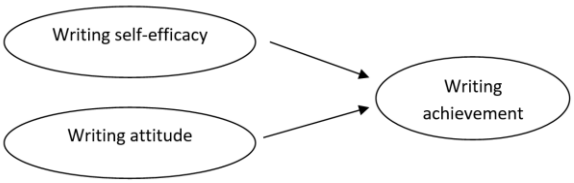


Figure 1. Writing Model

Methodology

Research Model

The current study was designed in the relational survey model, as it aimed to determine the relationship between the primary school fourth grade students' writing self-efficacy, writing attitude and summary writing achievement. This relationship was investigated by using the structural equation modeling. "Relational survey models are the models aiming to determine the existence of covariance between two or more variables and/or its degree." (Karasar, 2000, p. 81).

Participants

The universe of the study consists of the primary school fourth grade students in the city of Yozgat. In order to determine the sampling, the schools with medium socio-economic level from among the secondary schools in the city were determined and two of them were randomly selected to make up the sampling. A total of 335 students participated in the study (Girls: 160, Boys: 175).

Process

During the data collection process of the study, the students were asked to complete two scales (writing attitude scale, writing self-efficacy scale) and to write the summaries of two narrative texts. In the selection of the narrative texts, the following procedure was followed: First 8 narrative texts from two textbooks decided to be used with 4th graders by the Ministry of National Education were selected. As narrative texts are easier to summarize than informative texts (Anderson et. al 1991), the study only focused on narrative texts. Then the selected texts were sent to five experts working in the field of Turkish and language teaching. The experts examined the texts in terms of their suitability for the level of students. As a result of this examination, only three texts on which all the experts agreed were selected. Then the researcher selected two of them and asked the students to summarize them. During the summary writing process, the texts were given to the students for them to see and look at the texts as long as they wanted. Nothing more was said to the students. The students were given 40 minutes for each summary writing. The students were clearly explained how to fill in the scales on sample questions not included in the scales.

Measure

The data of the study were collected by using the Writing Self-efficacy Scale, Writing Attitude Scale, Summary Evaluation Form for Narrative Texts and Personal Information Form. The results obtained from the reliability and validity studies of the scales are given below:

Writing Self-efficacy Scale. The Writing Self-Efficacy Scale was developed by Güneş, Kuşdemir and Bulut (2017) in order to determine the writing self-efficacy status of primary school fourth grade students. The scale was prepared in the form of a three-point Likert scale (suitable for me, a little suitable for me, not suitable for me). However, considering the age and grade level of the students, next to these three response options, facial expressions reflecting their meanings were also placed to be able to obtain more reliable and valid data. The lowest score to be taken from the Writing Self-efficacy Scale is 10 and the highest score is 30. During the development of the scale, exploratory and confirmatory factor analyses were conducted to establish the validity of the scale. As a result of exploratory factor analysis, the three factors of the scale were found to be explaining 52.8% of the total variance and this is an acceptable level. For the whole of the scale, Cronbach Alpha value was calculated to be .69. Before and after the rotation, the total value of the variance explained by the factors did not change. As a result of confirmatory factor analysis, this three-factor construct was found to have adequate goodness-of-fit indices *(CMIN/DF = 2.14; GFI= .964; AGFI= .938; CFI= .914; RMR= .014; RMSEA= .057).* The Writing Self-efficacy Scale was administered to a group having similar characteristics to the ones possessed by the study group before it was used for the current study and the Cronbach Alpha value was calculated to be .72 for the whole scale on the collected data from this piloting application.

Summary Evaluation Form for Narrative Texts. In the current study, "The Summary Evaluation Form for Narrative Texts" developed by Bulut (2103) to evaluate students' summaries written for narrative texts was used. The Summary Evaluation Form consists of 23 items and 4 sub-dimensions. These sub-dimensions are content, form, text structure and language and expression. In the measurement of each criterion in the form, a three-point scaling was used as adequate = 2 points, partially adequate = 1 point and inadequate = 0 point. In order to establish the content validity of the Summary Evaluation Form, the researchers sought expert opinions and in addition to this, used Lawshe technique for the content validity. In line with this technique, for each item in the Summary Evaluation Form, the content validity ratio (KGO) was calculated. And for each item, a value higher than 0.62 was found (for 10 experts). Moreover, the content validity ratio calculated for the Summary Evaluation Form for Narrative Texts is 0.93. Therefore, all the items were included in the form. For the reliability of the Summary Evaluation Form, 50 summaries were evaluated by the researcher and by another researcher and the correlation coefficient between the results of the two researchers was calculated to be r= 0.82. The high and significantly positive correlation between these two evaluations indicates that the evaluation form yields reliable measurements.

Data Analysis

The data obtained from the administration of the Writing Self-efficacy and Attitude Scales and the summary texts written by the students were scored and the obtained results were entered into the computer environment. The collected data were analyzed by using SPSS 22.0 and AMOS 22.0 statistical programs.

Results

In the current study, the effect of writing self-efficacy and writing attitude on students' writing achievement was examined. In this section, findings obtained in the current study are presented. First, the descriptive statistics and correlations related to the variables of the study are given.

Table 1. *Correlation results related to the variables*

	Writing attitude	Writing self-efficacy	Summary writing score
Correlation (r)			
Writing attitude	1	.87	.72
Writing self-efficacy	.87	1	.53
Summary writing score	.72	.53	1

As can be seen in Table 1, there is a positive and very high correlation between writing attitude and writing self-efficacy, and between writing self-efficacy and summary writing. Moreover, there is a positive and high correlation between writing self-efficacy and summary writing.

In the study the hypothesis "writing attitude and writing self-efficacy significantly and positively affect summary writing achievement" was tested. In order to determine whether there is such an effect, the structural equation modeling was used. As a result of the analysis, the fit values were found to be as follows: (χ^2/sd= 2.15; RMSEA= 0.043; GFI= 0.90; AGFI= 0.90; CFI= 0.92). These values show that there is a good and acceptable fit. Thus, the hypothesis of the study was supported. That is, the effect of writing attitude and self-efficacy on primary school fourth grade students' summary writing achievement is significant and positive.

Discussion

Self efficacy is a concept that is influential on social and motor skills and affecting students' learning and achievement performance (Schunk, 1989). When an individual is asked to perform tasks for which his/her motivation is low, the concept of self-efficacy comes to the fore as a critical concept. Writing is one of these tasks (Bruning, Dempsey, Kauffman, McKim & Zumbrunn, 2013). Writing is a complex and demanding task that is posing motivational challenges for even talented writers. Conceptually, self-efficacy refers to some certain abilities. Writing requires many skills, strategies and knowledge such as self-regulation. Thus, for some certain aspects of writing, self-efficacy may change (MacArthur, Philippakos & Graham, 2015).

Writing is a difficult task for people from any age group (Troia&Graham, 2003) because it is not just a mechanical task but also a mental activity. Particularly students at school do not like writing or experience difficulty in writing tasks. This leads to the formation of negative attitudes towards writing and decreasing self-efficacy belief.

From among the affective factors influential on the writing skill, the most important ones are self-efficacy and attitude. When the relevant literature is reviewed, it is seen that there are some studies investigating the effect of students' writing self-efficacy and attitude on writing achievement (Pajares & Valiante, 1997; Graham, Berninger & Fan, 2007; Demir, 2013; Baştuğ, 2015). Karadağ and Kayabaşı (2013) conducted a study on pre-service teachers and found that the main reasons for their reluctance to write are lack of self-confidence (40.2%) and lack of motivation (16.1%).

Summary writing is a process of finding the main idea on the basis of the source text, removing unnecessary details and shortening by using one's own words. However, while doing this, the structure and flow of thoughts in the source text should not be distorted. As a result of the current study exploring the effect of writing self-efficacy and attitude on writing achievement through the structural equation modeling, the established hypothesis was supported. That is, the effect of writing self-efficacy and attitude on writing achievement is positive and significant. When the findings of the current study are compared to the findings reported by similar studies in the literature, they seem to support each other. Baştuğ (2015) conducted a study on writing and affective factors affecting it and found that writing disposition and attitude significantly and positively affect writing achievement; writing attitude and disposition significantly and negatively affect writer's block; writer's block significantly and negatively affect writing achievement. The greatest effect on writing achievement comes from writer's block, writing disposition and writing attitude, respectively. In this study, a high correlation was found between writing self-efficacy and writing attitude. Demir (2013) found a positive, moderate and statistically significant correlation between the primary school 8th graders' creative writing and self-efficacy perception. Moreover, a positive, moderate and significant correlation was found between the students' free writing skills and self-efficacy perception. A positive, low and significant correlation was found between the primary school 8th graders' topic-based written expression scores and self-efficacy perception. This finding indicates that the students' topic-based expression does not directly affect their writing self-efficacy perception. A positive, moderate and significant correlation was found between the primary school 8th grade students' creative writing scores and the writing self-efficacy scale's sub-dimension of written expression skills. Considering this significant correlation, it can be argued that the creative writing skills of the students whose written expression skill perceptions are high are developed. The research findings of Bruning et al. (2013) show that writing ideation and self-regulation self-efficacy to be significantly more strongly related to liking writing than conventions self-efficacy but less related than conventions self-efficacy to writing assessment scores. All 3 writing self-efficacy dimensions showed moderate positive correlations with self-reported writing performance.

In the current research, while examining the effect of writing self-efficacy and writing attitude on summary writing achievement, the students in the study group were asked to write the summaries of only narrative texts. Further research can be designed to explore narrative and informative texts together. Moreover, the study group of the current research consists of students from medium socio-economic level. Similar research can be conducted on students attending schools with different socio-economic levels.

References

Akar, C. (2008). Öz yeterlik inancı ve ilkokuma yazmaya etkisi. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, 1(2), 185-198.

Akaydın, Ş., & Kurnaz, H. (2015). Lise öğrencilerine yönelik yazma tutum ölçeği: geçerlik ve güvenirlik çalışması. *Mustafa Kemal University Journal of Graduate School of Social Sciences*, 12(32), 246-261.

Anderson, V., Suzanne, H., & Babadoğan, C. (1991). Özetlemenin öğrencilere öğretimi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 24(2), 587-594.

Baştuğ, M. (2015). Effects of primary school fourth-grade students' attitude, disposition and writer's block on writing success. *Education and Science*, 40 (180), 73-88.

Brown, A.L., Day, J.D., & Jones, R.S. (1983). The development of plans for summarizing texts. *Child Development,* 54, 968-979.

Bruning, R. , Dempsey, M., Kauffman, D.F., McKim, C., & Zumbrunn, S.(2013). Examining dimensions of self-efficacy for writing. *Journal of Educational Psychology*, 105(1), 25–38.

Bulut, P. (2013). *İlköğretim 5. sınıf öğrencileri ve öğretmenlerinin özetleme stratejilerinin değerlendirilmesi: Türkçe dersi örneği [*The evaluation of summarizing strategies of the primary school 5th grade students and teachers: Turkish lesson example *].* Unpublished Doctoral Dissertation, Gazi Üniversitesi, Ankara.

Can, B. (2016). *Öz düzenlemeli strateji geliştirmeye dayalı öğretim modelinin hikâye yazma becerilerini geliştirmeye etkisi. [* The effect of education model based on self-regulated strategy development on improving story writing skills]. Unpublished Doctoral Dissertation, Gazi Üniversitesi, Ankara.

Demir, T. (2013). İlköğretim öğrencilerinin yaratıcı yazma becerileri ile yazma özyeterlik algısı ilişkisi üzerine bir çalışma. *International Journal of Turkish Literature Culture Education,* 2(1), 84-114.

Göçer A. (2014). Yazma tutum ölçeğinin (ytö) Türkçeye uyarlanması: Geçerlik ve güvenirlik çalışması. *Kastamonu Üniversitesi Kastamonu Eğitim Dergisi*, 22(2), 515-524.

Graham, S., Berninger, V., & Fan, W. (2007). The structural relationship between wrting attitude and writing achivement in first and third grade students. *Contemporary Educational Psychollogy,* 32, 516-536.

Graham, S., Hebert, M., Sandbank, M.P., & Harris, K.R. (2016). Assessing the writing achievement of young struggling writers: Application of generalizability theory. *Learning Disability Quarterly,* 39 (2), 72-82.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology*, 99(3), 445–476.

Güneş, F. (2007). *Türkçe öğretimi ve zihinsel yapılandırma.* Ankara: Nobel Yayın Dağıtım.

Güneş, F., Kuşdemir, Y., & Bulut, P. (2017). Psychometric features of the writingself-efficacy scale. *International Journal of Social Science*, 58 ,101-114.

Jinks, J., & Lorsbach, A. (2003). Introduction: Motivation and self-efficacy belief. *Reading & Writing Quarterly,* 19(2), 113-118.

Karadağ, R., & Kayabaşı, B. (2013). Neden yazı yazmıyoruz? Sınıf öğretmeni adaylarının yazmayı engelleyen etmenlere ilişkin görüşleri. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi,* 15(1), 1-32.

Karasar, N. (2000). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.

MacArthur,C.A., Philippakos, Z.A., & Graham, S. (2015). Multicomponent measure of writing motivation with basic college writers. *Learning Disability Quarterly*, 39(1), 31-43.

Pajares, F. & Valiante, G. (1997). Predictive utility and causal influence of the writing self-efficacy beliefs of elementary students. Paper presented at the annual meeting of the American Educational Research Association. New York, US. Retrived from https://files.eric.ed.gov/fulltext/ED394144.pdf

Plotnik, R. (2009). *Psikolojiye giriş*. İstanbul: Kaknüs Yayınları.

Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review,* 1(3), 173-208.

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and selfevaluation. *Reading and Writing Quarterly*, 19, 159-172.

Senemoğlu, N. (2005). *Gelişim, öğrenme ve öğretim. Kuramdan uygulamaya*. Ankara: Gazi Kitabevi.

Troia, G. A., & Graham, S. (2003). The consultant's corner: "Effective writing ınstruction across the grades: what every educational consultant should know". *Journal of Educational and Psychological Consultation*, 14(1), 75-89.

Ungan, S. (2007). Yazma becerisinin geliştirilmesi ve önemi. *Erciyes Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 23 , 461-472.

Williams, A. R. (2007). *The effects of summarization ınstruction on the comrehension and metacognitive abilities of sixth grade students in a title I middle school*. Ph.D. Thesis. University of Alabama, Alabama.

Yıldız, M. ve Kaman, Ş. (2016). İlköğretim (2-6. sınıf) öğrencilerinin okuma ve yazma tutumlarının incelenmesi. *Türkiye Sosyal Araştırmalar Dergisi*, 2, 507-522.

This page is intentionally left blank

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

Developing a Teachers' Gender Stereotype Scale toward Mathematics*

Özge Nurlu [a]

Received: 14 August 2017
Revised: 06 Sept. 2017
Accepted: 18 Sept. 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236124

Abstract

Gender has become a focus of mathematics education research. While some research show that there are no differences between boys and girls, numerous research studies have indicated that boys have outperformed girls. It is suggested that gender stereotypes, such as expecting girls to show less achievement in mathematics compared to boys, have an effect on mathematics achievement. According to these gender stereotypes, boys are more successful in mathematics and science and girls are more successful in literature and arts. Gender stereotypes are transmitted by one generation to the next generation via children's books, language, parents and teachers as well. Because of teachers' important role of shaping their students' beliefs and attitudes, determining teachers' gender stereotypes is vital to understanding the differences of mathematical achievement between girls and boys. Therefore, the purpose of this study is to develop a teachers' gender stereotype scale toward mathematics. The scale consists of two subscales: the Boys' Form and the Girls' Form. These two forms are conducted with 595 primary school teachers. Results of the exploratory factor analysis for each form, 17 items and four factors are determined. Based on the literature review, these factors are named as environment, gender appropriateness of careers, competence and attribution of success. For each form, the confirmatory factor analysis is conducted and the four factors of the subscales are confirmed. The findings of the study revealed that the scale is a valid and reliable instrument to measure gender stereotypes in mathematics.

Keywords: Primary school education, mathematics, gender stereotype.

Introduction

Gender issues have become a focus of mathematics education research. Although some research shows that there are no differences between girls and boys (Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Kazu, & Ersözlü, 2008; Dede & Dursun, 2008; Yücel & Koç, 2011), a significant amount of research indicates that boys are more competent than girls in terms of cognitive variables such as problem solving and mathematical thinking abilities (Geary, Saults, Liu, & Hoard, 2000; Gallagher, De Lisi, Holst, McGillicuddy-De Lisi, Morely, M., & Cahalan, 2000; Altunçekiç, Yaman, & Koray, 2005) and affective variables such as mathematics anxiety, mathematical attitudes and self-efficacy (Köğce, Yıldız, Aydın, & Altındağ, 2009; Frenzel, & Pekrun, 2007; Kargar, Tarmizi, & Bayat, 2010; Çakıroğlu & Işıksal, 2009). In addition, some research investigated the mathematical achievement differences of girls and boys. Most of them find that boys are more successful than girls in mathematics (Van de Gaer, Pustjens, Van Damme, & De Munter, 2008; Tate, 1997).

Considering the results of research that attempted to find the relationship between gender and mathematical achievement, it is possible to wonder what kind of reasons could be effective. According to Weissglass (2002), several factors can affect students' mathematical achievement such as ethnicity, socio-economic status, language, sexual orientation, gender, the role of school, and culture as well. Researchers have conducted studies to investigate gender stereotypes in mathematics education as a part of culture (Spencer, Steele, & Quinn, 1999; Schmader, 2002; Brown, & Josephs, 1999; Schmader, Johns, & Barquissau, 2004). These gender stereotypes are the kind of beliefs that boys are more competent than girls in mathematics and science, and girls are more competent than boys in literature and arts (Beilock, Gunderson, Ramirez, & Levine, 2010). Studies find that these gender stereotypes are transmitted from one generation to the next generation via children's books (Taylor, 2003), language (Wigboldus, Sermin, & Spears, 2000), parents (Eccles & Jacob, 1986) and teachers (Esen, 2013; Keller, 2001).

Teachers' beliefs about mathematics have an effect on students' beliefs and even achievements (Beilock et al., 2010). Similarly, teachers' beliefs about mathematics as a male domain influence their students' beliefs and achievement in mathematics (Keller, 2001). Therefore, measuring teachers' gender stereotype beliefs toward mathematics is important to preventing the reproduction of gender stereotypes in mathematics in the classroom and providing a more balanced mathematics education environment for both genders. Even though there are various gender stereotype scale studies developed by

* This paper is part of an ongoing PhD dissertation project and supervised by Prof. Dr. Naciye AKSOY. This paper was an oral presented at the XV. USOS.

[a] Gazi University, Gazi Faculty of Education, Beşevler, 06500, Ankara, Turkey. Email: ozgenurlu@hotmail.com (Ö. Nurlu)

different researchers in the literature (Leder & Forgasz, 2002; Keller, 2001; Yee & Eccles, 1988; Tiedemann, 2000; Räty, Vänskä, Kasanen, & Kärkkäinen, 2002), these scales about gender stereotypes in mathematics are generally developed toward students and parents. Nevertheless, there are some research focus on measuring teachers' gender stereotypes in mathematics (Tiedemann, 2002; Keller, 2001). However, these studies use biased scales that provide participants with an opportunity to display only the degree of perceived masculinity of mathematics, and do not allow them to rate it as a female domain. For instance, participants who take a low score from a biased scale means that they have a low stereotypical belief about masculinity of mathematics. However, there is no evidence about the stereotypical belief about mathematics as a female domain. In this case, participants could have neutral beliefs in terms of gender in mathematics or even they could regard mathematics as a female domain. The new unbiased scale, included two subscales offering participants to indicate their beliefs about mathematics both as a male and female domain, is thought to be helpful for researchers who want to measure teachers' gender stereotypes in mathematics.

Methodology

Participants

The study is conducted with 595 primary school teachers in Turkey. It is considered that the teachers are experienced in teaching. 76% of the participants are female, while 24% of them are male.

Developing the Scale

The Teachers' Gender Stereotype Scale toward Mathematics is a scale applied as a just one scale however it consists of two subscales: Boys' Form and Girls' Form. During the first stage of the development process, literature is reviewed to determine categories of gender stereotypes in mathematics.

Table 1. *Mathematics Gender Stereotypes Indicators and Items*

Categorises	Indicators	Definitions	Sources	Item Examples
Attribution	Effort, ability, chance, support of parents or teachers, easiness of exam	It examines that teachers' attributions about the reasons and sources of students' achievement.	Yee, and Eccles (1988), Tiedemann (2000)	Compared to boys, girls mostly increase their mathematical achievement, because of the support of their teachers. Compared to girls, boys mostly increase their mathematics scores when the examination is too easy.
Competence	Having mathematical, logical thinking, problem solving abilities, motivation and confidence, discovering problems situations, searching patterns, verification of results, making generalisation	It examines that teachers' beliefs about students' mathematical knowledge, ability and attitudes.	Abrantes (2001), Tiedemann (2000), Leder, and Forgasz (2002), Fennema, and Sherman (1976) Milli Eğitim Bakanlığı (MEB) (2015)	Girls use mathematical tools such as rulers, number blocks etc., more effectively than boys do. Boys are more successful than girls in mental computation.
Effort	Contributing classroom activities, studying, seeking help, helping others, completed work in the classroom, making unassigned practice	It examines that teachers beliefs about students' effort in mathematics.	Brookhart (1997)	Boys complete tasks in mathematics classes more than girls do. Girls bring mathematics problems to ask their teachers more than boys do.
Career	Interest of jobs needed mathematical ability, characteristics of planned career, appropriateness of career	It examines that teachers' beliefs about students' career choice.	Correll (2001), Dick, and Rallis (1991), Rallis, and Ahren (1986)	Boys are encouraged more than girls to choose a career in a mathematically-related area.
Environment	Perceptions of peers and parents	It examines that teachers beliefs about how peers and parents perceptions on students in mathematics.	Parsons, Adler, and Kaczala, (1982), Andre, Whigham, Hendrickson, and Chambers (1999), Leder, and Forgasz (2002)	Compared to boys, girls are seen as more competent in mathematics by their parents.

In this regard, categories are written on the basis of scales developed by Leder & Forgasz (2002), Keller (2001), Yee and Eccles (1988), Tiedemann (2000) and Räty et al. (2002). According to these categories, indicators, definitions and items emerge. Table 1 shows these items based on the definitions. A pool with 42 items is written by considering each revised indicator related with gender stereotypes in mathematics. 6 of them take place in the attribution factor, 19 of them are in the competence factor, 6 of them are in the effort factor, 5 of them are in the career factor and lastly 6 of them are in the environment factor. The 42-item form that emerged is

analysed and evaluated by four experts from Primary School Education, six experts from Elementary School Mathematics Education, and one expert from Division of Curriculum and Instruction. Experts are asked to evaluate these items according to appropriateness in terms of ability to measure the gender stereotype beliefs, and intelligibility of items. Also, experts are asked to give suggestions if an item is inappropriate. According to feedback given by experts, intelligibility of some items is improved and 7 items are added for the competence factor. As a result, a 49-item form emerges. Items are written by giving superiority for each gender. For example, 'Boys are more competent than girls in using a calculator' and 'Girls are more competent than boys in using a calculator'. In order to determine participants' gender stereotype beliefs in mathematics, a 5-point Likert-type form is used.

Data Collection

The scale is first applied to 245 primary school teachers for explanatory factor analysis (EFA). After that, for confirmatory factor analysis (CFA), 350 primary school teachers complete the scale. The data collection process takes almost 6 months.

Data Analysis

Although some researchers suggest different sample size requirements to perform validity and reliability analysis, it is acceptable to reach 5-10 times the number of items on the scale (Kass, & Tinsley, 1979; Kline, 1994; Pett, Lackey, & Sullivan, 2003; Tavşancıl, 2005). Considering this criterion, 595 teachers are reached to fill out the scale. Before the start of the data analysis, extreme, outlier, and missing values are corrected. At the end of this, validity and reliability analysis are performed as a result of the answers gained from 595 primary school teachers. In this study, data are subsequently collected. In order to reveal the structure of the scale, EFA is performed on the first group (n_1= 245). CFA is performed on the other group (n_2= 350) to test the structure.

Findings

Findings Related to Validity

Before starting to define the structure of the Teachers' Gender Stereotype Scale toward Mathematics by performing EFA, in order to determine the aptitude of data gathered for Boys' Form and Girls' Form of the scale, the KMO and Bartlett's. Test are calculated. According to Kaiser (1974), a KMO value greater than 0.5 can be accepted. Pallant (2001) suggests that the KMO value should be higher than 0.6 to perform EFA. In this study, while the KMO value of Boys' Form is found as .90, the value of Girls' Form is calculated as .91. In this regard, the KMO values of these two forms are both greater than the values to be recommended by researchers. Bartlett's Test needs to have a significant value to determine the factorability of the correlation matrix obtained from the items. Bartlett's Test is found to be significant for both forms of the scale: Boys' Form χ^2= 2193.501; p= 0.00 and Girls' Form χ^2= 1863.416; p= 0.00. Therefore, it is possible to indicate that data from the trial form of the scale are proper for performing a factor analysis.

The total variance values of the items in the Boys' Form and Girls' Form are examined. It is seen in Figure 1a that items of Boys' Form are gathered under 4 factors which are bigger than 1.00 eigenvalues.

As it is clear from the Figure 1b, items of Girls' Form are actually gathered under 3 factors according to criterion of eigenvalue bigger than 1.00. Because of one more factor which is 0.905 eigenvalue has an important contribution to the scale, this factor is included as well. It is claimed that items with a factor loading above 0.4 are included in the output while items with a factor loading less than 0.4 need to be removed. Boys' Form factor loadings and variance values are seen in Table 2a.

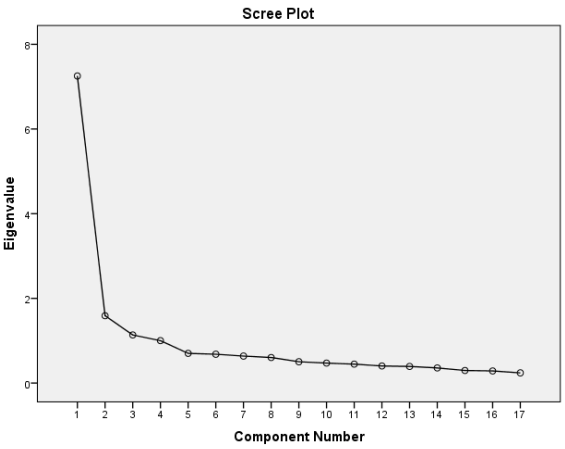


Figure 1a

The first extended factor consisted of 4 items ranging from .50 to .71, the second extended factor consisted of 4 items ranging from .70 to .74, the third extended factor consisted of 6 items ranging from .69 to .75 and the last extended factor consisted of 3 items ranging .64 to .78. Whole factors explain 64.5% of total variance. The first factor explains 26.75% of total variance and is labelled as 'environment'. The second factor explains 15.96% of total variance and is labelled as 'career'. The third factor explains 14.19% of total variance and is labelled as 'competence'. The fourth factor explains 9.67% and is labelled as 'attribution'.

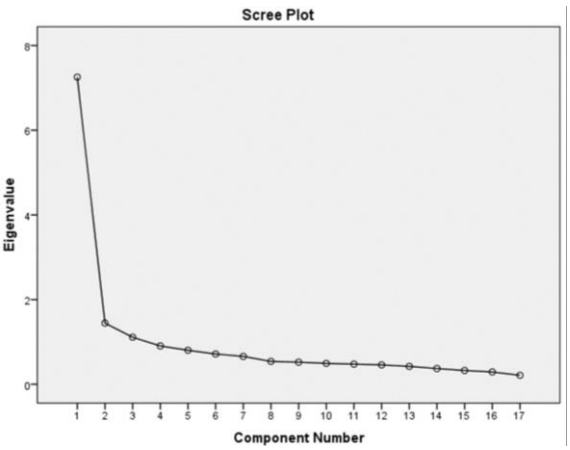


Figure 1b

Table 2a. *Teachers' Gender Stereotype Scale toward Mathematics: Boys' Form Factor Loads and Common Factor Variances*

	Items	Factor 1	Factor 2	Factor 3	Factor 4	Common Factor Variance
Environment	Compare to girls, boys are seen more competent in mathematics by their parents.	.565				.541
	Boys' parents think that mathematics is important more than girls' parents do.	.817				.705
	Compared to girls, boys are more popular because of their mathematical success.	.506				.555
	Boys are expected more than girls to do well in mathematics by their parents.	.714				.692
Career	Boys are more interested in careers which require mathematical ability than girls are.		.732			.731
	Boys are encouraged more than girls to choose a career in a mathematically-related area.		.727			.715
	Boys are more suited than girls to work in engineering branches.		.709			.686
	Boys are more willing than girls to work in mathematically-related areas.		.745			.756
Competence	Boys understand mathematical concepts more easily than girls do.			.692		.525
	Boys are more successful than girls in mental computation.			.701		.662
	Boys are more likely than girls to believe they can be successful in mathematics.			.733		.595
	Boys have higher logical thinking abilities than girls have.			.734		.617
	Boys have higher mathematical thinking abilities than girls have.			.752		.658
	Boys understand mathematical problems more easily than girls do.			.698		.664
Attribution	Compared to girls, boys mostly increase their mathematical achievement, because of the support of their teachers.				.789	.672
	Compared to girls, boys mostly increase their mathematics scores when the examination is too easy.				.604	.585
	Compared to girls, boys mostly increase their mathematics scores because their parents provide them with mathematical support.				.641	.622
	Eigen Value	7.252	1.590	1.136	1.001	
	Explained Variance	24.756	15.962	14.190	9.677	
	Total Explained Variance	64.586				

* Values lower than .40 is not shown in the table.

Table 2b. *Teachers' Gender Stereotype Scale towards Mathematics: Girls' Form Factor Loads and Common Factor Variances*

	Items	Factor 1	Factor 2	Factor 3	Factor 4	Common Factor Variance
Environment	Compared to boys, girls are seen as more competent in mathematics by their parents.	.574				.553
	Girls' parents think that mathematics is important more than boys' parents do.	.781				.677
	Girls are expected more than boys to do well in mathematics by their parents.	.750				.752
Career	Girls are encouraged more than boys to choose a career in a mathematically-related area.		.600			.629
	Girls are more suited than boys to work in engineering branches.		.747			.697
	Girls are more willing than boys to work in mathematically-related areas.		.795			.780
Competence	Girls are more successful than boys in predicting how to solve mathematical problems.			.657		.600
	Girls are more likely than boys to believe they can be successful in mathematics.			.632		.476
	Girls like solving mathematics problems that their classmates are not able to more than boys do.			.660		.527
	Girls are more successful than boys in describing the situation given in mathematical problems with mathematical symbols.			.819		.733
	Girls use mathematical tools such as rulers, number blocks etc., more effectively than boys do.			.801		.695
	Girls are more successful than boys in using a calculator in mathematics classes.			.669		.548
	Girls have higher mathematical thinking abilities than boys have.			.694		.591
	Girls are more successful than boys in modelling mathematical relationships by drawings.			.675		.560
Attribution	Compared to boys, girls mostly increase their mathematical achievement, because of the support of their teachers.				.782	.645
	Compared to boys, girls mostly increase their mathematics scores when the examination is too easy.				.757	.636
	Compared to boys, girls mostly increase their mathematics scores because their parents provide them with mathematical support.				.633	.616
	Eigen Value	7.254	1.445	1.111	0.905	
	Explained Variance	42.668	8.500	6.534	5.323	
	Total Explained Variance	63.026				

* Values lower than .40 is not shown in the table.

In Table 2b, Girls' Form factor loadings and variance values can be seen. According to Table 2b, first extended factor consisted of 3 items ranging from .57 to .78, the second extended factor consisted of 3 items ranging from .60 to .79, the third extended factor consisted of 8 items ranging from .63 to .81 and the last extended factor consisted of 3 items ranging from .63 to .78. Whole factors explain 63% of total variance. The first factor explains 46.66% of total variance and is labelled as 'environment'. The second factor explains 8.50% of total variance and is labelled as 'career'. The third factor explains 6.53% of total variance and is labelled as 'competence'. The fourth factor explains 5.32% and is labelled as 'attribution'.

For each form of the scale, correlations between factors are tested. Correlation coefficients between factors of Boys' Form are shown in Table 3a, of Girls' Form are shown in 3b.

Table 3a. *Correlation Coefficients between Factors of Boys' Form*

Factors	Environment	Career	Competence	Attribution
Environment	1.00	.599	.624	-.495
Career		1.00	.631	-.463
Competence			1.00	-.375
Attribution				1.00

**p<0.01

Table 3b. *Correlation Coefficients between Factors of Girls' Form*

Factors	Environment	Career	Competence	Attribution
Environment	1.00	.563	.538	-.480
Career		1.00	.585	-.434
Competence			1.00	-.447
Attribution				1.00

**p<0.01

As seen in Table 3a, correlation coefficients between factors of Boys' Form are ranging from -.37 to .62 and they are significant at .01 level.

As seen in Table 3b, correlation coefficients between factors of Girls' Form are ranging from -.43 to .58 and they are significant at .01 level.

CFA is performed to confirm the structure of the model revealed after EFA. χ^2/df chi-square/degree of freedom, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Goodness of Fit Index (GFI), Adjusted Goodness of Fit Index (AGFI), Normed Fit Index (NFI) and Comparative Fit Index (CFI) are taken into consideration as model fit indices. The four-factor model fit indices are determined as: Boys' Form: $\chi 2/df = 3.34$ (p= .000); RMSEA= .081; GFI= .90; AGFI= .84; CFI= .91; NFI= .88; SRMR= .06. Boys' Form is presented in Fig. 2a and 3a.

According to model fit indices, χ^2/df value for Boys' Form is 3.34, for Girls' Form it is 2.03. Kline (2005) states that there is a perfect match in models if the value is less than 2.5 for small samples. However, there is no consensus regarding for χ^2/df value. As Wheaton, Muthen, Alwin, & Summers (1977) indicate that less than 5.0 is acceptable ratio for this statistics. Therefore, these values in both Boys' and Girls' Forms are acceptable. The RMSEA value is found to be .08 for Boys' Form and .05 for Girls' Form. According to the literature, these values indicate a good cohesiveness (Brown, 2006). Additionally, GFI and AGFI values above .90 mean the model has perfect fit, and AGFI value above .80 is considered adequate (Jöreskog & Söbom, 1993). In this regard, GFI values are perfect and AGFI values are acceptable for both forms of the scale. As Sümer (2001) states that there is a good model fitting if CFI and NFI values are above .90. However, according to some researchers above .80 is acceptable, as well (Hair, Black, Babin & Anderson, 2009). Therefore, these values for the both forms of the scale are acceptable.

Cronbach's alpha is used as an estimate of the reliability of the scale. Cronbach's alpha values of Boys' Form are presented in Table 4a. Cronbach's alpha values of Girls' Form are presented in Table 4b.

Table 4a. *Cronbach's alpha values of Teachers' Gender Stereotype Scale towards Mathematics: Boys' Form*

Factors	Cronbach's Alpha Values
Environment	.771
Careers	.768
Competences	.915
Attribution	.580
General	**.884**

Table 4b. *Cronbach's alpha values of Teachers' Gender Stereotype Scale towards Mathematics: Girls' Form*

Factors	Cronbach's Alpha Values
Environment	.673
Careers	.593
Competences	.863
Attribution	.729
General	**.910**

According to Nunally (1978) alpha values higher than .70 are considered adequate. Cronbach's alpha value of Boys' Form is calculated as .884, and of Girls' Form is calculated as .910. In this regard, it is possible to indicate that both forms of the scale have adequate reliability.

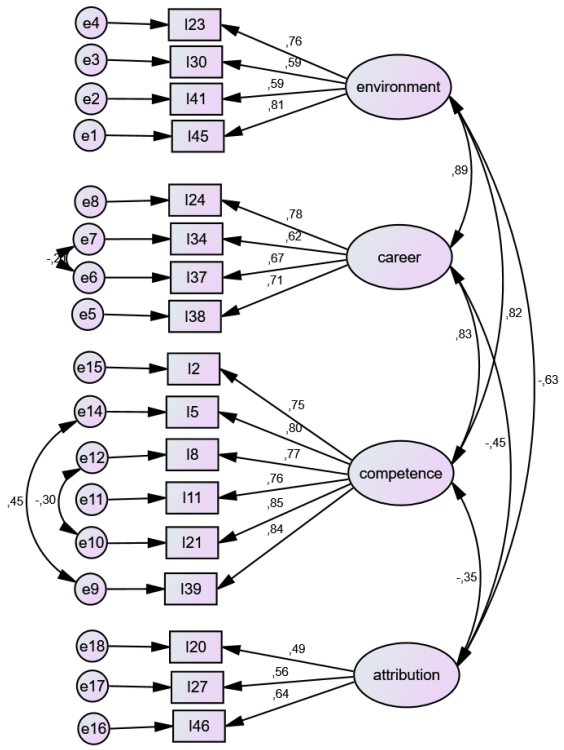


Figure 2a

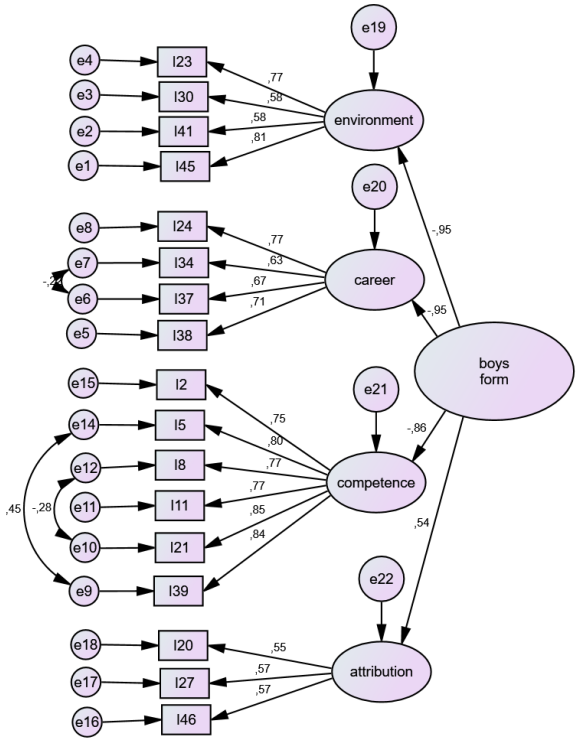


Figure 3a

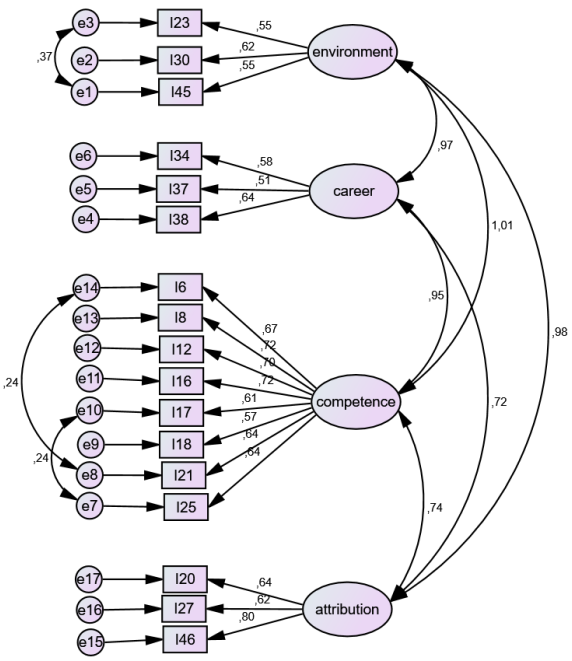


Figure 2b

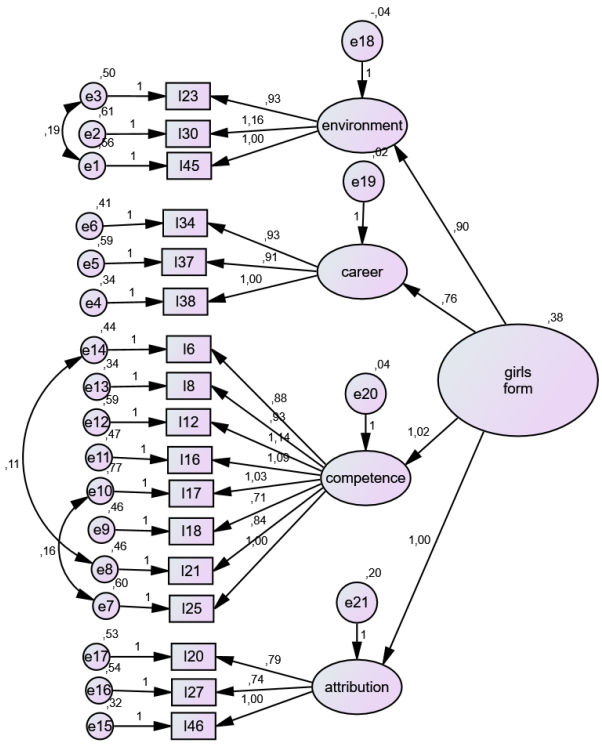


Figure 3b

Girls' Form: χ2/df= 2.03 (p= .000); RMSEA= .05; GFI= .93; AGFI= .90; CFI= .94; NFI= .90; SRMR= .04., Girls' Form is presented in Figure 2b and 3b.

Results

In this study, a new scale is developed to measure teachers' gender stereotype beliefs toward mathematics by considering mathematics gender stereotype indicators prepared by many researchers generally from western culture (Leder, and Forgasz, 2002; Keller, 2001; Yee, and Eccles, 1988; Tiedemann, 2000; Räty, et al. 2002). The scale has two subscales: Boys' Form and Girls' Form. Also it consists of four factors environment, career, attribution, and competence. However, it is possible to say that the literature has not offered a consistent structure about gender stereotypes about mathematics. Also, these studies are generally conducted to investigate parents' or children's beliefs about mathematics gender stereotypes. Nevertheless, there is a small amount of research investigating gender stereotype beliefs about mathematics particularly in teachers (Tiedemann, 2000, 2002) and these studies have limited sub-dimensions compared the other scales. This study focuses on teachers' beliefs and uses four scale factors existing in other research studies. Therefore, it is possible to say that, the research has an important role to investigate teachers' gender stereotype beliefs in mathematics field more comprehensively. According to the results, the scale is reliable and valid. In future, studies aimed to investigate teachers' mathematics gender stereotypes can use this scale.

References

Abrantes, P. (2001). Mathematical competence for all: Options, implications and obstacles. *Educational Studies in Mathematics, 47*(2), 125-143.

Altunçekiç, A., Yaman, S., & Koray, Ö. (2005). Öğretmen adaylarının öz-yeterlik inanç düzeyleri ve problem çözme becerileri üzerine bir araştırma (Kastamonu İli Örneği). *Kastamonu Eğitim Dergisi*, *13*(1), 93-102.

Andre, T., Whigham, M., Hendrickson, A. and Chambers, S. (1999). Competency beliefs, positive affect, and gender stereotypes of elementary students and their parents about science versus other school subjects. *J. Res. Sci. Teach.*, *36*(6), 719–747.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Brookhart , S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement . *Applied Measurement in Education, 10*(2), 161 - 180.

Brown, R. P., & Josephs, R. A. (1999). A burden of proof: stereotype relevance and gender differences in math performance. *Journal of Personality and Social Psychology*, *76*(2), 246.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford.

Correll, J. C. (2001). Gender and the career choice process: the role of biased self-assessments,A*merican Journal of Sociology*, *106*(6), 1691-1730.

Çakıroğlu, E., & Işıksal, M. (2009). Pre-service elementary teachers' attitudes and self-efficacy beliefs toward mathematics. *Education and Science, 34*(151), 132-139.

Dede, Y., & Dursun, Ş. (2008). İlköğretim II. kademe öğrencilerinin matematik kaygı düzeylerinin incelenmesi. *Eğitim Fakültesi Dergisi*, *11*(2), 295-312.

Dick, T. P., & Rallis, S. P. (1991). Factors and influences on high school students' career choices. *Journal for Research in Mathematics Education*, *22*(4), 281-292.

Eccles (Parsons), J., Meece, J. L., Adler, T. F., & Kaczala, C. M. (1982). Sex differences in attributions and learned helplessness. *Sex Roles*, *8*(4), 421-432.

Eccles, J. S. (1987). Gender roles and women's achievement-related decisions. *Psychology of Women Quarterly*, *11*(2), 135-172.

Eccles, J. S., & Jacobs, J. E. (1986). Social forces shape math attitudes and performance. *Signs, 11*(2), 367-380.

Esen, Y. (2013). Gender discrimination in educational processes: An analysis on the experiences of studentship. *International Online Journal of Educational Sciences, 5*(3), 757-782.

Fennama, E., & Sherman, J. A. (1976). Fennema-Sherman mathematics attitude scales: Instruments designed to measure attitudes toward the learning of mathematics by females and males. *Journal for Research in Mathematics Education*, *7*(5), 324-326.

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics*, *21*(1), 55-69.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Girls and mathematics—A "Hopeless" issue? A control-value approach to gender differences in emotions towards mathematics. *European Journal of Psychology of Education*, *22*(4), 497-514.

Gallagher, A. M., De Lisi, R., Holst, P. C., McGillicuddy-De Lisi, A. V., Morely, M., & Cahalan, C. (2000). Gender differences in advanced mathematical problem solving. *Journal of Experimental Child Psychology*, *75*(3), 165-190.

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology*, *77*(4), 337-353.

Hair J, Anderson RE, Tatham RL, Black WC. (2009). *Multivariate Data Analysis*. 7th ed. Upper Saddle River, New Jersey, Pearson Prentice Hall, USA.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science*, *321*(5888), 494-495.

Jöreskog, K., & Sörbom, D. (1993). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago, IL: Scientific Software International Inc.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika,* 39, 31-36.

Kargar, M., Tarmizi, R. A., & Bayat, S. (2010). Relationship between mathematical thinking, mathematics anxiety and mathematics attitudes among university students. *Procedia-Social and Behavioral Sciences*, *8*(2010), 537-542.

Kass, R. A., & Tinsley, H. E. A. (1979). Factor analysis. *Journal of Leisure Research*, 11, 120-138.

Kazu, H., & Ersözlü, Z. N. (2008). Öğretmen adaylarının problem çözme becerilerinin cinsiyet, bölüm ve öss puan türüne göre incelenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, *8*(1), 161-172

Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology*, *141*(2), 165-173.

Kline, P. (1994). *An Easy Guide to Factor Analysis*. New York: Routledge.

Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling* (2nd ed.). New York: Guilford Press.

Köğce, D., Yıldız, C., Aydın, M., & Altındağ, R. (2009). Examining elementary school students' attitudes towards mathematics in terms of some variables. *Procedia-Social and Behavioral Sciences*, *1*(1), 291-295.

Leder, G. C., & Forgasz, H. J. (2002). Two new instruments to probe attitudes about gender and mathematics.

MEB, (2015). Milli Eğitim Bakanlığı talim ve terbiye kurulu başkanlığı ilköğretim matematik dersi (1-4.sınıflar) öğretim programı.

Nunnally JC. (1978). *Psychometric Theory*, 2nd ed. New York: McGraw-Hill.

Pallant, J. (2001). *SPSS Survival Manual*. Maidenhead: Open University Press.

Parsons, J. E., Adler, T. F., & Kaczala, C. M. (1982). Socialization of achievement attitudes and beliefs: Parental influences. *Child Development*, *53*(2), 310-321.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making Sense of Factor Analysis: The Use of Factor Analysis for Instrument Development in Health Care Research*. CA: Sage.

Rallis, S. P., &Ahren, S. A. (1986). *Math and Science Education in High Schools: A Question of Sex Equality?* American Educational Research Association, San Francisco, California

Räty, H., Vänskä, J., Kasanen, K., & Kärkkäinen, R. (2002). Parents' explanations of their child's performance in mathematics and reading: A replication and extension of Yee and Eccles. *Sex Roles*, *46*(3-4), 121-128.

Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, *38*(2), 194-201.

Schmader, T., Johns, M., & Barquissau, M. (2004). The costs of accepting gender differences: The role of stereotype endorsement in women's experience in the math domain. *Sex Roles*, *50*(11-12), 835-850.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, *35*(1), 4-28.

Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları*, *3*(6), 49-74.

Tate, W. F. (1997). Race-ethnicity, SES, gender, and language proficiency trends in mathematics achievement: An update. *Journal for Research in Mathematics Education*, *28*(6), 652-679.

Tavşancıl, E. (2005). *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. Ankara: Nobel.

Taylor, F. (2003). Content analysis and gender stereotypes in children's books. *Teaching Sociology*, *31*(3), 300-311.

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school Mathematics. *Educational Studies in Mathematics*, *41*(2), 191-207.

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, *50*(1), 49-62.

Van de Gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2008). Mathematics participation and mathematics achievement across secondary school: The role of gender. *Sex Roles*, 59(7-8), 568-585.

Weissglass, J. (2002). Inequity in mathematics education: Questions for educators. *The Mathematics Educator, 12*(2), 34-39.

Wigboldus, D. H., Semin, G. R., & Spears, R. (2000). How do we communicate stereotypes? Linguistic bases and inferential consequences. *Journal of Personality and Social Psychology*, *78*(1), 5.

Wheaton, B., Muthen, B., Alwin, D., F., & Summers, G. (1977). Assessing Reliability and Stability in Panel Models. *Sociological Methodology, 8*(1), 84-136.

Yee, D. K., & Eccles, J. S. (1988). Parent perceptions and attributions for children's math achievement. *Sex Roles*, *19*(5-6), 317-333.

Yücel, Z., & Koç, M. (2011). The relationship between the prediction level of elementary school students' math achievement by their math attitudes and gender. *Elementary Education Online*, *10*(1), 133-143.

Appendix

1.Items in Turkish Language

	Hiç Katılmıyorum	Katılmıyorum	Biraz Katılıyorum	Katılıyorum	Tamamen Katılıyorum
1. Kızlar erkeklere göre velileri tarafından matematikte daha yeterli görülürler.					
2. Erkekler kızlara göre matematik dersinde iyi oldukları için arkadaşları arasında daha popülerdirler.					
3. Kızlar erkeklere göre velilerinde daha yüksek matematiksel başarı beklentisi oluştururlar.					
4. Erkekler kızlara göre aileleri tarafından matematik ile alakalı bir iş koluna yönelmeleri konusunda daha çok desteklenirler.					
5. Kızlar erkeklere göre çeşitli mühendislik alanlarında çalışmaya daha uygundurlar.					
6. Erkekler kızlara göre velileri tarafından matematikte daha yeterli görülürler.					
7. Erkekler kızlara göre matematik başarılarını daha çok, sınav kolay olduğunda arttırabilirler.					
8. Kızlar erkeklere göre matematik dersinde sınıf arkadaşlarının çözemediği soruları yanıtlamaktan daha çok hoşlanırlar.					
9. Kızlar erkeklere göre problem sonuçlarını tahmin etmede daha başarılıdırlar.					
10. Kızlar erkeklere göre matematik dersinde araç ve gereçleri (cetvel, onluk sayı blokları vb.) daha etkili bir şekilde kullanırlar.					
11. Erkekler kızlara göre matematik başarılarını daha çok, velileri matematiksel destek sağladığı için arttırırlar.					
12. Kızlar erkeklere göre daha üst düzey matematiksel düşünme becerilerine sahiptirler.					
13. Kızlar erkeklere göre matematik başarılarını daha çok, sınav kolay olduğunda arttırabilirler.					
14. Erkekler kızlara göre matematikle alakalı bir iş kolunda çalışmayı daha çok isterler.					
15. Erkekler kızlara göre daha üst düzey matematiksel düşünme becerilerine sahiptirler.					
16. Erkekler kızlara göre velilerinde daha yüksek matematiksel başarı beklentisi oluştururlar.					
17. Kızlar erkeklere göre aileleri tarafından matematik ile alakalı bir iş koluna yönelmeleri konusunda daha çok desteklenirler.					

	Hiç Katılmıyorum	Katılmıyorum	Biraz Katılıyorum	Katılıyorum	Tamamen Katılıyorum
18. Kızlar erkeklere göre matematik başarılarını daha çok, velileri matematiksel destek sağladığı için arttırırlar.					
19. Kızlar erkeklere göre daha çok, matematiğin en önemli ders olduğunu düşünen velilere sahiptirler.					
20. Erkekler kızlara göre matematik kavramlarını daha kolay anlarlar.					
21. Kızlar erkeklere göre matematik problemlerinde verilen bir durumu matematiksel sembollerle göstermede daha başarılıdırlar.					
22. Erkekler kızlara göre daha çok, matematiğin en önemli ders olduğunu düşünen velilere sahiptirler.					
23. Kızlar erkeklere göre matematikle alakalı bir iş kolunda çalışmayı daha çok isterler.					
24. Erkekler kızlara göre çeşitli mühendislik alanlarında çalışmaya daha uygundurlar.					
25. Kızlar erkeklere göre matematik başarılarını daha çok, öğretmen onlar ile ilgilendirdiği için arttırırlar.					
26. Erkekler kızlara göre matematik problemlerini daha kolay anlarlar.					
27. Erkekler kızlara göre zihinden işlem yapmada daha başarılıdırlar.					
28. Kızlar erkeklere göre matematikte başarılı olabileceklerine daha çok inanırlar.					
29. Erkekler kızlara göre matematik başarılarını daha çok, öğretmen onlar ile ilgilendiği için arttırırlar.					
30. Erkekler kızlara göre daha üst düzey mantıksal düşünme becerilerine sahiptirler.					
31. Kızlar erkeklere göre matematik dersinde hesap makinesi kullanmakta daha başarılıdırlar.					
32. Erkekler kızlara göre matematikte başarılı olabileceklerine daha çok inanırlar.					
33. Erkekler kızlara göre üst düzey matematiksel beceri gerektiren meslek kollarına daha çok ilgi duyarlar.					
34. Kızlar erkeklere göre matematiksel ilişkileri resimler yoluyla modellemede daha başarılıdırlar.					

2. Items in English Language

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. Compared to boys, girls are seen as more competent in mathematics by their parents.					
2. Compared to girls, boys are more popular because of their mathematical success.					
3. Girls are expected more than boys to do well in mathematics by their parents.					
4. Boys are encouraged more than girls to choose a career in a mathematically-related area.					
5. Girls are more suited than boys to work in engineering branches.					
6. Compare to girls, boys are seen more competent in mathematics by their parents.					
7. Compared to girls, boys mostly increase their mathematics scores when the examination is too easy.					
8. Girls like solving mathematics problems that their classmates are not able to more than boys do.					
9. Girls are more successful than boys in predicting how to solve mathematical problems.					
10. Girls use mathematical tools such as rulers, number blocks etc., more effectively than boys do.					
11. Compared to girls, boys mostly increase their mathematics scores because their parents provide them with mathematical support.					
12. Girls have higher mathematical thinking abilities than boys have.					
13. Compared to boys, girls mostly increase their mathematics scores when the examination is too easy.					
14. Boys are more willing than girls to work in mathematically-related areas.					
15. Boys have higher mathematical thinking abilities than girls have.					
16. Boys are expected more than girls to do well in mathematics by their parents.					

	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
17. Girls are encouraged more than boys to choose a career in a mathematically-related area.					
18. Compared to boys, girls mostly increase their mathematics scores because their parents provide them with mathematical support.					
19. Girls' parents think that mathematics is important more than boys' parents do.					
20. Boys understand mathematical concepts more easily than girls do.					
21. Girls are more successful than boys in describing the situation given in mathematical problems with mathematical symbols.					
22. Boys' parents think that mathematics is important more than girls' parents do.					
23. Girls are more willing than boys to work in mathematically-related areas.					
24. Boys are more suited than girls to work in engineering branches.					
25. Compared to boys, girls mostly increase their mathematical achievement, because of the support of their teachers.					
26. Boys understand mathematical problems more easily than girls do.					
27. Boys are more successful than girls in mental computation.					
28. Girls are more likely than boys to believe they can be successful in mathematics.					
29. Compared to girls, boys mostly increase their mathematical achievement, because of the support of their teachers.					
30. Boys have higher logical thinking abilities than girls have.					
31. Girls are more successful than boys in using a calculator in mathematics classes.					
32. Boys are more likely than girls to believe they can be successful in mathematics.					
33. Boys are more interested in careers which require mathematical ability than girls are.					

This page is intentionally left blank

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

The Effects of The Living Together Through Art (LTTA) Model on Promoting The Consciousness of Living Together Between Thai and Migrant Students

Received: 14 June 2017
Revised: 31 August 2017
Accepted: 17 Sept. 2017
ISSN: 1307-9298

www.iejee.com

Sarita Juaseekoon[b] , Apichart Pholprasert[a,*], Chanita Rukspollmuang[c]

DOI: 10.26822/iejee.2017236125

Abstract

The aim of this research is to study the effects of the Living Together through Art model (LTTA model), a newly developed art learning model based on the concept of UNESCO's "Learning to Live Together", on promoting the consciousness of living together between Thai and migrant students which consisted of 4 core values: respect, acceptance, empathy and appreciate. The principle of the LTTA model was "Using art learning activities to encourage the ethnically-mixed students to express themselves, connect and collaborate with each other". The research was carried out using quasi-experimental methods. The trial was conducted at a primary school in Samut Sakhon, Thailand in the 2016 academic year. Forty-one ethnically mixed students form second and third grade classrooms were divided into an experimental group and a control group. Data collection consisted of a test, behavior observation, students' reflection and in-depth interviews. The data were analyzed using means, standard deviation, one-way ANCOVA, repeated ANOVA and content analysis. The data revealed that (1) the posttest mean score of the experimental group was higher than the pretest mean score and the score of the control group at a significant difference of .05 which remained stable after 4 weeks and (2) the frequency of desirable behavior which related to the core values increased and (3) students expressed more positive thoughts about themselves and others.

Keywords: Learning to live together, migrant students, prejudice reduction, consciousness, art learning.

Introduction

Population mobility, including international labor migration, has been on the increase in recent decades. This phenomenon leads to cultural diversity which presents challenges for any educational system. Schools must uphold the right to equal education for every learner and support their needs, but also promote an understanding of cultural diversity among students so they can live with each other in harmony, especially, in the context of migrant inclusion that could cause tensions between majority and minority groups.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) attempts to foster the understanding, tolerance and friendship between youth of all nations, as well as racial or religious groups which is necessary for the maintenance of peace. In 1996, the concept of "Learning to Live Together" (LTLT) was originally set out in a report for UNESCO by the International Commission on Education for the Twenty-First Century chaired by Jacques Delors as one of the 'Four Pillars of Education'. The report emphasized that the survival of humanity is highly dependent on learning how to live together, beginning by understanding and accepting other people and their history, cultures, traditions and values. (Delors et al., 1996).

Delors (1996) stated that LTLT results from two complementary processes: the "discovery of others" and the "experience of shared purposes". "Discovery of others" means learning about self, others and society. Students have to realize that human beings are the same because we are all human, but different because we are culturally diverse but, we are all dependent on each other. This process will help children learn to respect, have empathy and accept others. 'Experience of shared purposes', by playing or working together towards a common goal, could change the potential tension between diverse group into friendship. If one has the opportunity to communicate with others, they will be able to understand and appreciate different points of views that may lead to prejudice reduction.

There is a relationship between LTLT process and the approaches of prejudice reduction from two social psychology theories: Tajfel and Turner's social identity theory (1979) and Allport's contact hypothesis (1954). Tajfel and Turner believed that prejudice is a result of

*,a Corresponding author: Apichart Pholprasert, Faculty of Education, Chulalongkorn University, Thailand. Phone: +668 4144 7004 E-Mail: apicharr@hotmail.com
[b] Sarita Juaseekoon. Chulalongkorn University, Thailand. E-mail: sarita.new@gmail.com
[c] Chanita Rukspollmuang, Siam University, Thailand. E-mail: drchanita@gmail.com

group membership. People maintain their self-esteem in part by identifying with groups and believing that the groups they belong to are better than other groups. In order to decrease prejudice, the differences within groups and the similarity between groups should be exaggerated, which relates to the "Discovery of others". Meanwhile, Allport suggested that interpersonal contact is one of the most effective ways to reduce prejudice between majority and minority group members. However, the positive effects of intergroup contact occur in contact situations characterized by four key conditions: equal status, intergroup cooperation, common goals, and support by social and institutional authorities. This concept is related to "Experience of shared purposes". These connections could explain the potential of LTLT to establish understanding and relationships among people from different cultures.

There have been and continue to be many educational initiatives designed to teach the concept of LTLT, such as peace education, multicultural education, human rights education. Regardless of name, all initiatives aim to change participants' internalizing skills, values and behavior. As Delors said about "creating a new spirit" (p. 22) which leads to new perception and action, students should be cultivated in the level of consciousness to make the change from inside to outside. Sinclair (2004) claimed that LTLT pedagogy should be active and constructive where students could explore and construct their own understanding as needed. Both cognitive and affective domains should be involved. Furthermore, interactive and collaborative activities should be included so students can learn from each other.

The potential of art in promoting LTLT

Art has great potential to foster the concepts of LTLT. As seen in UNESCO's report 'Learning to Live Together: Education Policies and Realities in the Asia-Pacific'(2006), 9 of 10 have chosen countries choose art education as non-academic carrier subjects to transfer LTLT competencies in students. Art is a good vehicle for LTLT learning because it is strongly bound with culture. Firstly, as Efland, Freedman and Stuhr (1996) have suggested, art is defined as "a form of cultural production" whose value lies in its ability to promote "deeper understandings of the social and cultural landscape" (p. 72). So, art can construct learners' objective knowledge about self, other and social environment. Secondly, art has a social significance. Vygotsky (1971) asserted that "Art is the social technique of emotion, a tool of society which brings the most intimate and personal aspects of our being into the circle of social life" (p. 249). Crossing the boundaries from one's experiences to others' leads to empathy and can develop the ethic of care. (Greene, 1995) Thirdly, as learning process, meanings creating and understanding art, is a transformative experience that brings students to see, experience, appreciate and value aspects of the world in new ways enabling the construction of their own ideas, making new meaning and expressing it through artwork. (Dewey, 1974; Eisner, 2002) Eisner (1972) noted that "the visual arts deal with an aspect of human consciousness that no other field touches on" (p. 9). Lastly, group activities in an art program, such as making mural or group discussion, can promote engagement and collaboration among students (Day and Hurwitz, 2012).

Previous research has used art as a tool to reduce the prejudice and build the positive relationship among students from different background which is related to the goal of LTLT. For example, the Arab-Jewish Class Exchange Program by Berger, Abu-raiya and Gelkopf (2015) integrated art learning with contact hypothesis theory to reduce stereotyping and prejudicial attitudes between Israeli-Jewish and Israeli-Palestinian children. The results show that this program aided students to increase their readiness for contact with students from other groups, to express more positive thoughts and exhibited less emotional prejudice. Another example is the Story Telling/ Mural Painting project by Kim and Wiehe-Beck (2016) that involved the collaborative processes of storytelling, writing and art to promote international understanding among fifth grade students of diverse ethnicities. It was found that the project succeeded in changing the attitudes of student participants 2/3 of whom responded that they wanted to change how they act toward others. These studies indicate acknowledge the potential of art to reduce prejudice and fostering positive perceptions in students. However, both studies integrated art learning with other activities, such as writing, music or drama. Consequently, the researcher intended to design an LTLT learning model using art (visual art) as a core element of the model that appropriate with the context of migrant inclusion in Thai local school.

The situation of migrant inclusion in Thai local school

Thailand has a long history of hosting migrants and refugees from neighboring countries as they flee conflict or poverty. From the beginning of the 1990's, the number of migrant workers, the majority form Myanmar, has increased steadily. The increased numbers of migrant workers has resulted in a higher number of migrant children in the country. Given the risk of human trafficking and reacting to the policy of Education for All (EFA), the Thai government decided to stipulate in law that all children, regardless of their nationality or legal status, have the right to 15 years of free basic education (Ministry of Education, 2016). This law provides the opportunity for migrant students to attend Thai public school. However, most Thai schools adhere to a curriculum which includes teaching more about national identity, 'Thainess', than to teaching about cultural diversity (Anantasuchartkul, 2011). Moreover, Thai perceptions of refugee and migrant workers tends to be rather negative because of the belief that migrant workers pose a threat to public safety, may carry diseases, compete for jobs with Thais and national resources (Sunpawan and Niyomsilpa, 2012). These misperceptions could lead to prejudices that can obstruct the development of relationships between Thai and migrant students. Therefore, it is essential that schools with migrant students should focus on prejudice reduction and promoting international understanding.

The researcher conducted at survey to examine the situation of migrant inclusion in 6 selected schools in Samut Sakhon using non-participant observation and interviews of the school's principals, teachers and

students and their parents, including NGOs and government officials. The results reveal that prejudice was found not only between Thai and migrant students, but among migrant students, too. Moreover, some migrant students, especially the second generation who were born in Thailand, tried to conceal their ethnicity and claimed to be Thai because they didn't want to be labeled "alien". Following from these challenges, the researcher aimed to develop an art learning model that can reduce prejudice among students and promote consciousness of living together, which suitable to the context of a culturally inclusive classroom in Thailand, so that both Thai and migrant students could develop a positive perception of themselves and each other.

The development of the Living Together through Art (LTTA) model

Living Together through Art (LTTA) is a newly developed model based on the concept of UNESCO's "Learning to Live Together". As this model aims to achieve prejudice reduction, the theoretical premise of contact hypothesis and social identity theory that related to Delors' LTLT complementary processes were used, and combined with the method for reducing classroom prejudice described by Cushner, McCleland and Safford (2006). Their study suggest that improving social contact and intergroup relations, increasing cognitive sophistication, improving self-confidence and self-acceptance and increasing empathy for and understanding of others is possible. In addition, from Banks and Banks' (2001) review of the literature about multicultural education, this model focused on the role of teachers and the learning environment as supporting factors.

The researcher combined the emphasis on prejudice reduction and art learning processes, and then synthesized the ideas into the model's principles. The principle of the LTTA model was "Using art learning activities to encourage the ethnically-mixed students to express themselves, connect and collaborate with each other". This principle consist of 5 components: 1) expressing self through art, 2) comparing aspects of sameness and differences between persons, 3) imagining from others' viewpoint, 4) learning from personal and cultural narratives, and 5) collaborating for shared goals. After that, the completed model (figure 1) was created.

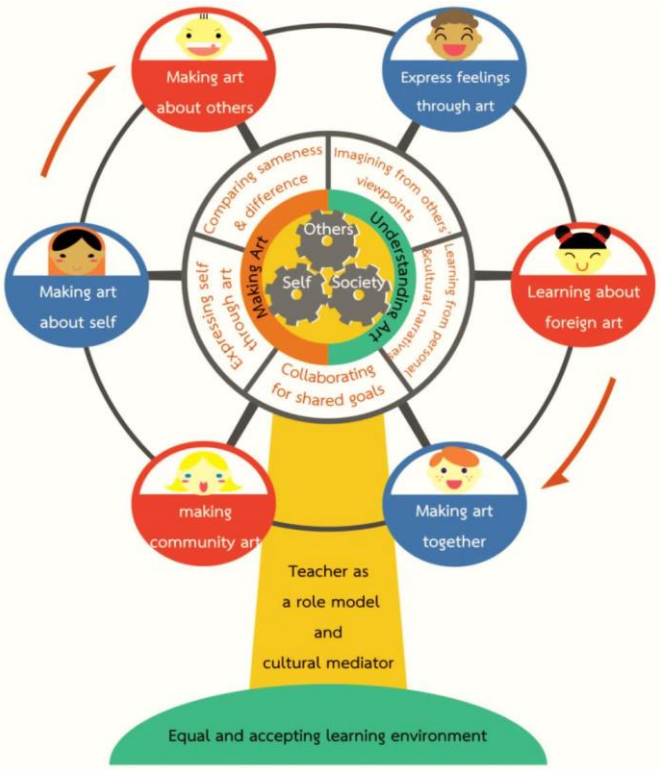


Figure 1. *Living Together through Art (LTTA) model*

The LTTA model used the figure of a Ferris wheel to represent a place where children of different ethnicities can play together. The core of the model consists of the main themes (self, others and society), an art learning process (making art and understanding art) and the 5 components as mentioned above. The components were linked to the learner in 6 activities; making art about self, making art about others, expressing feelings through art, learning about foreign art, making art together and making community art. The teacher's role and the proper learning environment were the supporting factors that facilitate the model run smoothly.

The model's activities were suggested briefly so teachers could adapt them to the needs of the students and school. The activities shown in table 1 are samples used in this study. Each session took 2 hours (the last session took 4 hours), that fit into the period of a school's extra-curricular activities.

Table 1. *LTTA Activities*

Session	Activities
1. It's me!	Students identify their characteristic and express through self-portrait. After, the class presents artworks and discusses identities.
2. Our future	Students learn about styles of portrait. Students work in pairs, sharing feelings about what they want to be when they grow up and drawing friends' portrait as imagined in the future.
3. The saddest day	Students identify feelings in famous artworks. They reflect on the saddest experience in their life, and express their feelings through art.
4. Travel through art	Groups of students compare the sameness and differences of their art from foreign art, e.g. comparing Thai, Mon and Indian pagodas.
5. Our colors	Students learn about the symbolic meanings of colors in their ethnic flags and use these colors to make paper quilts. Then, they arrange all the quilts to make a quilt mural.
6. We love our school	Students explore around their school and interview the principal, teachers and friends about the school. Students share what they like about school. Then, students design and make a school mural together.

The learning process in each activity was developed from a learning process of Arigatou Foundation's learning to live together program (2008); motivation, exploration, dialogue, discovery, reflection and action, and the meaning making in the art making process from Walker (2001); exploring the 'big idea', personal connection, building a knowledge base, problem solving, setting boundaries and designing studio instructions. The LTTA learning processes consisted of perceiving, connecting, creating, presenting, discussing and reflecting. The 'perceiving' phase began with artwork that related to the theme of the activities. Games and stories were used to stimulate the participants' curiosity about the artwork and topic. In the 'connecting' phase, the teacher would present facts while using questions to motivate students to think and connect the idea discussed with their own knowledge and experiences to construct new meaning. In the 'creating' phase, students would create artwork to express their meaning, before presenting it to a classmate in the 'presenting' phase. In the 'discussion' phase, there was an opportunity to exchange ideas, share experiences and discover those of others, which may challenge their perception and lead to new realizations. After that, in the 'reflecting' phase, participants took some time to revise their learning experience and write their reflection. This process proposed meaningful experience that led to individual consciousness transformation (Dewey, 1974; Eisner, 2002).

As the process of model development was completed, the study aimed to determine the effects of the LTTA model on students' consciousness of living together. For that purpose, the research questions addressed by this study included:

1. What is the effect of the LTTA model on participants' consciousness of living together?

2. What are the improvements in participants' perceptions and behaviors throughout the implementation process?

Method

Participants and Setting

The study was conducted at Samut Sakhon, Thailand. This province has a high concentration of migrant workers because of the demand from the fishery industry. Therefore, many Thai public schools in Samut Sakhon have students from migrant families. The research field is the small elementary school selected following the suggestion of 3 experts. The school has 174 students, 63.89% of whom are migrant students. Most were Mon and Burmese from Myanmar. Thai and migrant students study together in an inclusive class age and are ethnically mixed.

There were 41participants in this study from second and third grade in the academic year of 2016: 21 girls and 20 boys, aged 7-15 years. In regards to ethnicity, 53.66 percent were Mon, 24.39 percent Thai, 19.51 percent Burmese and 2.44 percent were Shan. All the migrant students could speak, read and write Thai. The participants were matched by sex, age and ethnicity and divided into two groups, 21 in the experimental group and 20 in the control group.

Data Collection Tools

As Farthing (1992) has stated "Consciousness is the subjective state of being currently aware of something either within oneself or outside of oneself" (p. 6) and "it (consciousness) concerns perception, thought, feelings and actions" (p.7). It could be stated that consciousness consists of inner awareness and the external behavior that expressed. Therefore, there were 2 groups of data collection tools. The test, interview and students' reflection were used to assess the inner awareness of consciousness while recorded behavior was used to examine external the action of consciousness. It was determined from a literature review and expert interview that the core values of the consciousness of living together that suited the Thai and migrant students'

context were respect, acceptance, empathy and appreciation.

Test of consciousness of living together

The test of consciousness of living together, developed by the researcher, was designed to take into consideration the outcome of the LTTA model. The test consisted of 25 questions, as statements or situations that related to all 4 core values that participants had to consider and make a decision about as to their level of agreement toward each question. A typical 5-level Likert scale was used and the following quantitative values were given: (1) strongly disagree, (2) disagree, (3) neither agree nor disagree, (4) agree, (5) strongly agree. The opinions and suggestions of 5 experts were sought to verify the reliability and validity of the questions. The reliability coefficient of these questions was administered to 35 third grade students who were familiar with the background information required for this unit. Using Cronbach's Alpha, the reliability coefficient was determined to be 0.806.

Behavior recording form

As mentioned, the change in the consciousness could be seen through behavior. This study used a behavior recording form to count the frequency of desirable behaviors (behaviors that related to the core values, e.g. showing the interest in other's artwork) and undesirable behaviors (behaviors that are opposite to the core values, e.g. paying no attention to friends' opinions) in each activity.

Students' reflection

Students were requested to write what they felt and thought after each activity. After this procedure, all the reflections were collected and used as a data source for researchers to understand what students had thought and learned.

Interview

At the beginning and the end of the procedures, a semi-structured interview was conducted to determine the students' perceptions about him/herself, others and society, especially viewpoints about people and classmates from different cultures. Each student was interviewed individually at an appropriate time. The interviews were audio-recorded.

Implementation Process

The study was carried out using a quasi-experimental method and according to the pre-test post-test design with non-equivalent group. Firstly, the participants were matched and divided into experimental and control groups, then the pre-test and the interview were conducted. Then, the experimental group was taught with the LTTA model, while the control group attended the school's extra-curricular activities. The implementation went on over four weeks at the rate of four hours a week, with a behavior recording and student reflection in every session. After that, the post-test and the interviews were conducted. Lastly, the experimental group did the follow-up test four weeks after the end of the implementation to examine the persistence of consciousness development.

Results

The results of the test of consciousness of living together

To examine whether there was any significant difference between the experimental group and control group regarding their level of consciousness of living together and whether the development of the consciousness of living together could remain over time, the test was measured 3 times: pre-test and post-test in both group and follow-up test in just the experimental group. The results are described below.

Table 2. *Means and standard deviations of experimental group's score of consciousness of living together test*

Test	*Experimental Group*		*Control Group*	
	Mean	*SD*	*Mean*	*SD*
Pre-test	84.71	7.90	84.60	9.01
Post-test	92.48	10.92	82.20	9.22
Follow-up test	94.67	11.18	-	-

A one-way ANCOVA was conducted to determine a statistically significant between the experimental group and control group regarding their level of consciousness of living together, controlling for pre-test score. Repeated measure ANOVA was also used to investigate the change in the experimental group's mean scores over three times.

From the data in table 2, a One-way ANCOVA results showed that there is a significant effects of the LTTA model on the score of the consciousness of living together after controlling for pre-test scores [F (1,38)= 12.58, p= .001]. The partial Eta Squared value was .249, comparing with Cohen's guidelines (Cohen, 1988), the effect size of the LTTA model was almost medium ($0.10 < \eta^2 \leq 0.25$). It can be said that LTTA Model was the factor that made a significant different on consciousness of living together.

Moreover, a repeated measure ANOVA with a Sphericity Assumed correction showed that the mean of the experimental groups' consciousness of living together score differed significantly between time points [F(2,40)= 11.399, p= .000]. Post hog tests using the Bonferroni correction revealed that the mean score increased by an average of 7.762 after the model implementation procedures (p= .003<.05) and remained the same after 4 weeks (p= 1.000>.05). In other words, it can be implied that after the LTTA model implementation, there was an improvement in students' consciousness that leveled off after 4 weeks.

Behavior Observation Results

To study the effects of the model on students' behavior, the participants' desirable and undesirable behaviors in each activity were counted. The means of behavior frequency are shown in the bar chart in figure 2.

The bar chart illustrates the mean of students' behavior frequency per session. It can be seen that the frequency of desirable behavior grew steadily and reached a peak in session 5, then dropped a little in session 6. While the mean of desirable behavior increased, the undesirable behavior decreased.

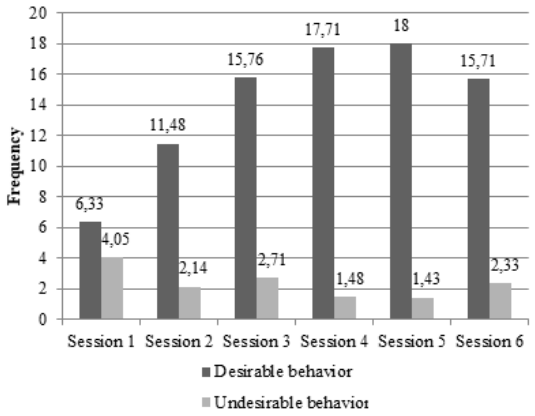


Figure 2. *The development of the experimental group's type of behavior within the procedure*

Even though the bars fluctuated over all session, the mean in session 6 was about 42% less than in session one. Based on these findings, it can be concluded that the art learning model was able to reinforce participants' desirable behavior.

Students' Reflections

Students' reflections expressed their feelings on what they learned from each activity. Some can be seen as evidence that students reached the objectives of the activities and the core values of the model. Students knew more about themselves, others, culture and diversity that led to more positive attitude and interest in their own and other culture. In addition, the reflections presented that besides learning from the content or activities, students learned from others' feelings and opinions which they felt related, so they could transform into the meaningful learning.

> *"At first, I saw myself in a negative way, but my friends helped me by telling me the things that they like about me. That's made me feel very happy. (It's me!)*
>
> *"If I make a mess on a friend's portrait, they may be upset and do the same to me. Therefore, I should treat others how I want to be treated. (Our future)*
>
> *"Today, my friend presented a picture about her bullying story that made me feel so sad. Next time, if I see someone being bullied, I will help them. (The saddest day)*
>
> *"Diversity is good. Like colors, various colors can make things colorful and beautiful!" (Our colors)*
>
> *"I like Thai puppets, but Myanmar puppets are beautiful as well. I think it can't be judged which one is better than the other, because they are both good." (Travel through art)*
>
> *"Today, we talked about what we like in our school. The conversation reminded me of the good moments. Even though it's small, I still love my school." (We love our school)*

Moreover, students also reflected their feeling that they like the activities in every session, because they were fun, attractive, challenging and differ from ordinary art class.

> *"I think the game of matching artwork to the artist is so much fun." (Our future)*
>
> *"Today we had to compare the sameness and difference of a piece of artwork from Thailand and other countries. The teacher let us use a magnifying glass to examine the pictures. I feel like I'm a detective."(Travel through art)*
>
> *"I am really interested in the story of national flags. Maybe tomorrow, I will search online for more information. (Our colors)*

Interviews

The interview was conducted before and after the implementation of the model. Before the start of the activities, some students expressed negative views toward those from different cultures, their own ethnicity or even the culturally inclusive classroom. The prejudice not just exists between Thai and migrant students, but among migrant students from different ethnicities, as a result of parental attitudes and their experiences.

> *"Some Thai classmates and seniors always call me 'Burmese', even though I'm not. Some of them make fun of my Thai accent. That's so annoying."*
>
> *"I speak Thai all the time, even in my house. My father didn't like that, but I don't want to be called 'Alien'. I wish I could be Thai so I can stay in Thailand forever."*
>
> *"I don't like the Burmese, they are cruel. My grand mom told me that they invaded and killed lots of Mon people. I don't want to be friends with them"*
>
> *"Dad told me not to play with Mon or Burmese. I don't like them, too. When they talk in their language that I can't understand, I feel like they are gossiping about me. I wish I could move to another schools that has only Thai students"*

Interestingly, after the implementation of the model which the students learned more about the positive aspects of diversity and worked collaboratively, students expressed more positive thought about self, others and their multicultural school. Moreover, some students realized that their prejudice beliefs were not true, so they decided to have a more open mind for other ethnic group.

> *"I'm glad to learn something about the Mon. When we're working in a group, I can tell others that 'this is a Mon flag' 'this is a Mon pagoda'. Some Thai friends asked me about the Mon language; that made me feel so proud of myself and my race."*
>
> *"I like the last activity, school murals. Everybody really put their minds to it! We did our best to make it beautiful. We want to show others that our school is good and unique, our school is multicultural!"*
>
> *"Previously, I didn't like the Burmese. But now I know that some of them are good. They are kind, bright and always take care of me and juniors in our group."*

Discussion

This study aimed to examine the effectiveness of the LTTA model to promoting the consciousness of living together between Thai and migrant students in an inclusive classroom. The results of the consciousness of living together test revealed that there was an increase in the test scores in students in the experimental group who participated in the LTTA model, as well as a significant difference between control and experimental group scores. In addition, the ANOVA results show that the increased in students' consciousness test scores did not

declined after 4 weeks, thus demonstrating that, their improvement in consciousness remained over time. With respect to the findings, it could be said that the LTTA model succeeded in promoting participants' consciousness of living together in a multi-ethnic community because it was effective in reducing students' prejudices and promoting core values. The effectiveness of the LTTA model could be the result of integrating prejudice reduction into the art learning process. In this study, students from diverse ethnicities were encouraged to learn about each other in both direct and indirect ways. When they understood and accepted other students, their attitude changed. (Berger, Abu-raiya and Gelkopf, 2015; Kim and Wiehe-Beck, 2016) Moreover, by increasing students' cognitive skills, self-confidence and empathy through activities, core values were constructed (Cushner, McCleland and Safford, 2006). Therefore, it could be claimed that the LTTA model is capable of reaching the goal of UNESCO's Learning to Live Together program.

Another interesting finding was that students did not change in their score alone, but their behavior and perception also developed in positive ways. Students' undesirable behaviors such as name calling, teasing, threatened and ridiculed were decreased. They showed more kindness and caring to each other. Most of students interested in others' culture, even someone who had expressed negative views toward other ethnic groups became more open-minded. According to Farthing (1992), consciousness concerns perceptions, thoughts, feelings and actions. It can be claimed that the students developed their level of consciousness with respect to students of other ethnicities and cultures. This finding could be a benefit of art for three reasons. First, art could actualize cultural concepts so that students could understand the others' cultural forms. As Efland, Freedman and Stuhr said that art is "a form of cultural production" (1996). Students may not understand about cultural diversity, but they could sense it when it was compared to the use of colors in artworks, for instance. Also, the teacher was able to bring students' cultural background into art class by using artwork that made the lesson about ethnicity easier to comprehend. Second, art could motivate students to focus on the topic intentionally, even topics they were not familiar with, such as ethnicity or empathy. Students had to reflect, connect with their experience, construct their own meaning and express that through art. Then, that meaning would be embedded in their mind with effects on their actions. As Eisner (2002) stated, art is how individuals interact with the world and leads to a 'complex and subtle form of thinking' that take place when children create meaningful artwork. The experience of meaning making leads to consciousness transformation. Third, the experience of the art learning process could be translated into their real life. For example, while students compared and contrasted a work of art, they were encouraged to compare their self to others effectively, learn that everybody shares a degree of sameness and difference manifest in artworks.

Furthermore, students' feedback revealed that they were really interested in LTTA activities. They claimed that LTTA activities were fun and different from ordinary art lessons, such as just drawing or painting. The themes of the artworks and art making studied were related to their daily life. These feelings led to students' mutual engagement, regardless of their culture or ethnicity. As Antoniou and Hickman (2012) claim, children's engagement in creating and responding to artworks will increased if the activities are meaningful, enjoyable and useful for them.

Moreover, activities that students' engaged in during the process of LTTA beyond art making, such as playing games, listening to stories, working in groups or discussing artwork, succeeded in motivating students. Integrating all these activities with art encouraged students to analyze facts, think critically, formulate question and collectively make decisions. Meanwhile, these could challenge and motivate students to think and make decision in collaboration with others, using and valuing the expertise of peers (La Porte, 2016). As the Arigatou foundation (2008) suggested in their LTLT program that the methodology of LTLT should place the learner in a self-driving learning process, conducted in relation to others, so each process should be designed to promote active participation, involvement and connection with others. Since the success of the LTTA model is due to the students' engagement, teachers should be concerned about designing activities related to students' interests and to motivate students whether individually or collectively.

Conclusions

This study showed the efficacy of LTTA model on promoting the consciousness of living together between Thai and migrant students in the context of culturally inclusive classroom. It was found that LTTA model that integrated the concept of UNESCO's "Learning to Live Together", prejudice reduction and multicultural education into the art learning process was succeed in cultivating students' consciousness of living together. In addition to enhanced students' understanding of cultural diversity, the qualitative data revealed the improvement in their attitude and behavior toward other ethnicities. Students demonstrated a positive change through the process of learning through art following the appreciation of an LTTA model. The results could be seen as the support of the notion that art is important for education, in a wide range of disciplines aside from visual art. It does not mean that art could improve students' academic performance, but art provides a meaningful experience for cultivation both self-consciousness and mutual understanding in a global community, to prepare students to meet any challenges in the modern world. Therefore, schools should apply some of LTTA model approaches to promote the consciousness of living together, regardless of whether schools are in the context of migrant inclusion.

Because the LTTA model was designed for use in Thai, it is not appropriate for application with migrant students who cannot communicate in Thai. In addition, due to the small of sample sizes in this study, these results may or may not be generalizable to other populations. Therefore, further studies with different classes of students in different contexts are required. In addition, it would be

beneficial for the LTTA model to be used outside of a school setting, such as among NGO, special education center or in community learning centers.

Acknowledgement

Sarita Juaseekoon was supported with a scholarship from "The 100th Anniversary Chulalongkorn University Fund for Doctoral Scholarship".

References

Allport, G. (1954). *The Nature of Prejudice.* Reading, MA: Allison-Wesley.

Anantasuchartkul, B. (2011). [Education Models for the Second Generation Migrants from Myanmar. Bangkok, Thailand: The Thailand Research Fund.

Anderson, T. (2003). Art Education for Life. *International Journal of Art and Design Education, 22*, 58-66.

Antoniou, P. and Hickman, R. (2012). Children's engagement with art: Three case studies. *International Journal of Education through Art*. 8(2), 169-182.

APCEIU. (2008). *Teacher Training for Learning to Live Together: A Training Manual for EIU and ESD.* 21 March 2016 retrieved from http://http://education4resilience.iiep.unesco.org/en/node/581

Arigatou Foundation. (2008). *Learning to Live Together: An Intercultural and Interfaith Programs for Ethics Education*. Geneva: ATAR Roto Press Sa.

Banks, J. A. and Banks, C. M. (2001). *Multicultural Education: Issue and Perspectives.* New York: John Wiley & Sons.

Berger R. and Abu-Raiya, H. (2015). The Art of Living Together: Reducing Stereotyping and Prejudicial Attitudes through the Arab-Jewish Class Exchange Program (CEP). *Journal of Educational Psychology*, *107*(3), 678-688.

Cohen, J. (1988). *Statistic Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cushner, K., McCleland, A. and Safford, P. (2006). *Human Diversity in Education.* NY: McGraw-Hill.

Day, M. and Hurwitz, A. (2012). *Children and Their Art: Art Education for Elementary and Middle Schools.* Boston: Wadsworth Cengage Learning.

Efland, A., Freedman, K., & Stuhr, P. (1996). *Postmodern art education: An approach to curriculum.* Reston, VA: The National Art Education Association.

Eisner, E.W. (1972). *Educating Artistic Vision*. London: Collier-Macmillan.

Eisner, E.W. (2002). *The Art and the Creation of Mind.* New Heaven & London: Yale University Press.

Farthing, G.W. (1992). *The Psychology of Consciousness.* Englewood Cliffs, N.J.: Prentice Hall.

Greene, M. (1995). *Releasing the Imagination: Essays on Education, the Arts and Social Change.* San Francisco: Jossey-Bass.

International Commission on Education for the Twenty-first Century., Delors, J., & UNESCO. (1996). *Learning, the treasure within: Report to UNESCO of the International Commission on Education for the Twenty-first Century*. Paris: UNESCO Pub.

Kim, J.H. and Wiehe-Beck, A. (2016). Understanding "the Other" through Art: Fostering Narrative Imagination in Elementary Students. *International Journal of Education & the Arts, 17*(2). 21 January 2017 retrieved from http://www.ijea.0rg/v17n2/.

La Porte, A.M. (2016). Efficacy of the Arts in a Transdisciplinary Learning Experience for Culturally Diverse Forth Graders. *International Electronic Journal of Elementary Education*, *8*(3), 467-480.

Ministry of Education (Thailand). (2016). [Manual and Guidelines for Non-Thai or Migrant Education]. Bangkok, Thailand: Thailand Agricultural Cooperative Press.

Sinclair, M. (2004). *Learning to Live together: Building Skills, Values and Attitudes for 21st Century*. Paris: The United Nations Educational, Scientific and Cultural Organization.

Stephan, W.G. (2014). Intergroup Anxiety: Theory, Research, and Practice. *Personality and Social Psychology Review*. *18*(3), 239-255.

Sunpawan, M. and Niyomsilpa, S. (2012). *Perception and Misperception: Thai Public Opinions on Refugees and Migrants from Myanmar.* Journal of Population and Social Studies, *21*(1), 47-58.

Tajfel, H. and Turner, J.C. (1979). An Integrative of Intergroup Conflict. In W. G. Austin, & S. Worchel (Eds.), *The social psychology of intergroup relations*. Monterey, CA: Brooks/Cole.

UNESCO. (2014a). *Learning to Live Together: Educational Policies and Realities in the Asia-Pacific.* Paris: UNESCO.

UNESCO. (2014b). *Teaching Respect for All: Implementation Guide.* 21 March 2016 retrieved from http://unesdoc.unesco.org/images/0022/002279/227983e.pdf

Wasson, R., Stuhr, P. and Petrovich-Mwaniki, L. (1990). Teaching Art in the Multicultural Classroom: Six Position Statements. *Studies in Art Education*, *31*(4), 234-246.

iejee INTERNATIONAL ELECTRONIC JOURNAL OF ELEMENTARY EDUCATION

December 2017, Volume 10, Issue 2

An Analysis of Pre-service Elementary School Teachers' Skills in Geometrical Drawing Using Isometric Paper

Veli Toptas[a,*], Elif Tugçe Karaca[a]

Received: 2 September 2017
Revised: 10 October 2017
Accepted: 18 October 2017
ISSN: 1307-9298

www.iejee.com

DOI: 10.26822/iejee.2017236126

Abstract

The aim of this study was to determine pre-service elementary school teachers' capabilities of deciding the viewpoint and perspective in geometrical drawing. The study examined geometrical drawings the participants did on isometric paper. This is a case study, a qualitative study method, and the study data were analyzed using written documents. The study sample included 34 senior students in the Education Faculty's Elementary School Teaching Program at a Turkish public university. They were selected from approximately 205 pre-service teachers. The students participated in the study on a voluntary basis. The study results revealed that the pre-service teachers generally had weak skills in geometrical drawing. The participants had better skills in building structures which they saw from the air using unit blocks, and forming a structure of which they had the direct views (left and right; front and back) using unit blocks. However, the participants had weaker skills in drawing the direct views of the structures they had built, recognizing the connections between the left-right views and front-back views of the structures and drawing the structures they built on isometric paper in perspective.

Keywords: Drawing skills, geometry, ısometric paper, pre-service primary school teacher, mathematics teaching.

Introduction

It has been suggested that the introduction to geometrical figures and naming, building, drawing, comparing and classifying them should be highlighted in the early years of primary education. This will also make a positive contribution to students' learning to analyze the shapes of the objects around them and recognize, name and draw their geometrical forms (Ministry of National Education, 2005).

According to the standards of the National Council of Teachers of Mathematics (NTCM, 2000), a mathematics curriculum should include the study of one, two and three-dimensional figures in a variety of situations so that the students will be able to determine, describe, compare and classify geometrical figures. Students learn about the geometrical figures by building, drawing, measuring, visualizing comparing, reshaping and classifying them. They also discover the correlations between them and develop their spatial intuition (Powel, 1997).

Geometrical drawing is a sub-field of mathematics, and its version in the elementary mathematics curriculum (MNE, 2009) includes activities that can help students develop spatial reasoning skills. Spatial reasoning is the depiction and understanding of the imaginary movements of two- and three-dimensional objects in space (Clements & Battista, 1992). According to NCTM (2000), school mathematics is supposed to improve students' spatial reasoning skills in order for them to comprehend space. The development of skills such as understanding space and drawing it, producing models of space and modifying them, and landscaping are based on geometrical thought. Individuals who have spatial reasoning skills are able to observe things in a more meaningful way and solve mathematical problems in daily life more easily (Duatepe & Paksu, 2013).

Geometrical drawing skill is important in teaching, learning and in the assessment of students' learning. An analysis of mathematics curricula indicate that both primary and elementary curricula include learning acquisitions in building three-dimensional structures using two-dimensional views and drawing three-dimensional views on isometric papers using perspective.

Moreover, NCTM standards stress the importance of the development of geometrical thinking and the drawing skills in the process of geometry teaching. Elementary school teaching programs' curricula in education faculties give importance to pre-service teachers' levels of three-dimensional thinking and drawing for both mathematics lessons and improving students' spatial thinking skills. However, the curricula do not include a course that focuses on these skills.

Since drawing is not about perception but representation, Piaget and Inhelder claimed that inaccurate drawings indicate the lack of mental tools that are required for spatial representation. Holloway (1967) said that a viewpoint and changes required by that viewpoint were necessary for an object to be represented in perspective by means of mental image or drawing (Toptaş, 2007).

[a,*] Corresponding author: Veli TOPTAŞ, Kirikkale University, Faculty of Education Department of Elementary Education, Kirikkale/Turkey. E-mail: vtoptas@gmail.com
[b] Kirikale University, Faculty of Education, Department of Elementary Education, Kirikkale/Turkey. E-mail: tugcekaraca85@gmail.com

Mitchelmore (1976, 1980) described children's drawings of spatial figures in four main stages. In the first stage, which is the schematic plane stage, drawings are two-dimensional or show only one face implying that the child looks at one face of the object in an orthogonal way. The second stage includes the schemas of concrete objects, and at this stage, children are able to draw pictures which can depict many faces, the seen and unseen faces of concrete objects, and they can draw pictures which can or cannot represent their depths. The third stage is the pre-realistic stage, which only includes drawings not having a clear or appropriate perspective of visual faces and depth. The fourth stage is the realistic stage, which involves drawing parallel lines at the correct lengths to depict parallelism and depth (Toptaş, 2007).

According to Hoffer (1981), drawing is one of the basic skills which should be taught to students in geometry. Hoffer claims that teaching drawing in geometry helps students to convey their thoughts using shapes. According to Piaget and Inhelder (1957), the unchanging qualities of geometrical shapes, the differences between shape-object relationships or Euclidian geometry depends on the relationships between different shapes or objects. The former is included in a particular shape and the latter covers the projective relations between shapes and objects or the Euclidian relationships among each other. Piaget and Inhelder assert that this type of a global coordination is a main necessity for the creation of a simple shape-object relationship from children's perspective. In the creation of this type of object-shape relationships which seem to be complicated, a single viewpoint is not formed separately but through the formation of a holistic system connection combined with other compulsory viewpoints.

Distinguishing three-dimensionality requires the use of perspective. According to Piaget (1970) and Morss (1987), the recognition and development of perspective is a great achievement. The representation of perspective requires a conscious coordination between the object and the subject. In other words, it is required to be known that object, subject and the observer are in the same plane depending on the viewpoint (Piaget and Inhelder, 1967). Children understand that objects look different from different perspectives when they are seven years-old, yet they cannot draw or imagine these changes in their viewpoints (Holloway, 1967). They can only understand how an object looks from a specific viewpoint when they see picture drawn from a particular viewpoint in perspective (Holloway, 1967). Children start to be able apply the rules of perspective systematically when they are eight or nine (cited by Authors, 2007).

It has been reported that children develop their geometrical thinking by discovering geometrical objects through physical contact, that they create perspective by drawing figures, and that touching objects had a positive effect on their spatial skills in mental rotation and spatial visualization (Clements & Battista, 1992; Werthessen, 1999).

For all these reasons, it is important that pre-service teachers' skills of perspective, viewpoint and drawing are determined and analyzed. Thus, this study aimed to determine pre-service teachers' ability levels in using perspective in geometrical drawing.

Methodology

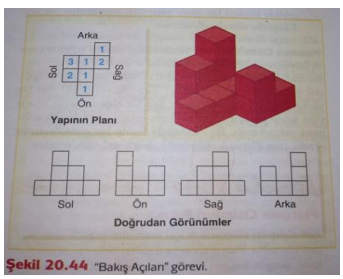
This is a case study, a qualitative study method, including the analysis of written documents. Document analysis can either be a research method itself or used as a supplementary information source when other qualitative methods are used (Yildirim & Simsek, 2011).

Participants

The participants of the study included 34 senior students in the Education Faculty's Elementary School Teaching Program in a mid-Anatolian University. As the design of the research is a case study, the participants were selected with convenience sampling method from 105 students in the program on a voluntary basis. 29 of the participants were female and 5 of them were male. The researcher also attended the participants' Mathematics Teaching 1 and 2 courses.

Instruments

The study examined the participants' skills in geometrical drawing using isometric paper. For this purpose two tasks entitled "Viewpoints" and "Perspective drawing" was used. The tasks were taken from "Geometrical thinking and concepts" chapter of Elementary and Middle School Mathematics: Teaching Developmentally Elementary and Middle School Mathematics by Van de Walle, Karp and Bay-Williams (2009). Students who enrolled the study took Mathematics Teaching Course-I and Mathematics Teaching Course-II. In those courses, Elementary and Middle School Mathematics: Teaching Developmentally Elementary and Middle School Mathematics by Van de Walle, Karp and Bay-Williams (2009) book was studied during the semesters. Also those tasks are appropriate for the participants' content knowledge level as they had been taught about 3D constructing and drawing 2D forms with isometric papers.

1a. 

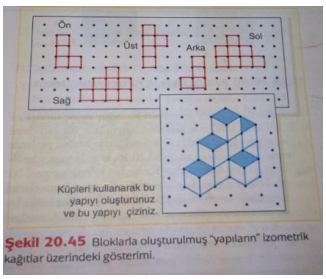
1b. 

Figure 1. *The activities used as data collection tools*

Procedure

In the first task, identifying a visualization task, one consideration is the degree of the attention that must be given to the particular properties of shapes. One of the main goals of the visualization strand is to be able to identify and draw two-dimensional images of three-dimensional figures and to build three-dimensional figures from two-dimensional images. Thus, the pre-service teachers were asked to draw the viewpoints and perspectival drawings shown in the figure. In the activity shown in Figure 1a, the students started by building a structure. The diagram of the structure shows an aerial view of it and the number of blocks in each position. Using this diagram, the students were asked to build the structure and draw views of its left, right, back and front faces.

In the perspectival drawing activity, the students were given a diagram that showed five direct views. Then, they were asked to build a structure according to this diagram. Participants began with an isometric 3D drawing of a building. From the drawing, participants were asked to build the actual building with their unit cubes. To record the result, they were required to draw a plan (top view) indicating the number of blocks in each position. After they built it, they were asked to draw the view of the right-front perspective. The two tasks were completed by the students about 45 minutes.

The authors used the textbook to analyze the participants' drawings and examine their geometrical drawing skills. The participants' ability levels in depicting geometrical viewpoints and perspective were analyzed regarding their geometrical drawing skills. With document analysis techniques the researchers first created a rubric for each task and used giving frequency and percentages about each task and sub-questions in each task.

Results

The authors interpreted the findings from these activities using frequencies and percentage tables.

Table 1. *Pre-service Teachers' Ability to Build The Structure With Unit Blocks*

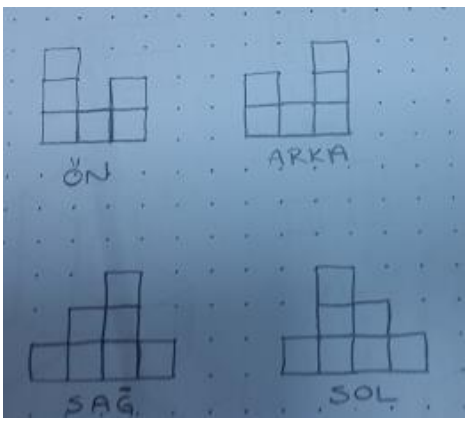
Ability to build the structure	f	%
Able to build the structure	32	94
Unable to build the structure	2	6
Total	34	100

Table 1 presents the findings about participants' ability to build a structure with unit blocks using an aerial view diagram. Of the participants, 94% were able to build the structure correctly using unit blocks. Of the participants, only 6%, or two participants, were unable to build the structure using the diagram. These findings indicated that pre-service teachers were highly capable of building a structure using unit blocks according to a given plan, and they were generally successful at this activity.

Table 2. *Pre-service teachers' ability to draw the views of the structure they built*

Views	f	%
Front	15	44 drew it correctly
Back	13	38 drew it correctly
Left	12	35 drew it correctly
Right	10	29 drew it correctly

Table 2 presents the findings about the participants' ability to draw views of the structure they built with unit blocks on isometric paper. The findings indicate that the participants' ability to draw views of the right, left, back and front faces of the structure was below 50% in general. The frontal view of the structure was drawn most successfully by 44% of the students. The least successful view of the structure was from the right with a failure rate of 71%. The achievement rates of the participants in drawing the structure from the back and left were 38% and 35%. In general, the participants' ability to draw the structures they built from a direct view was below 50%, which was lower than expected. Samples of correct drawings by the participants were shown in Figure 2, and samples of incorrect drawings are shown in Figure 3.

2a. 

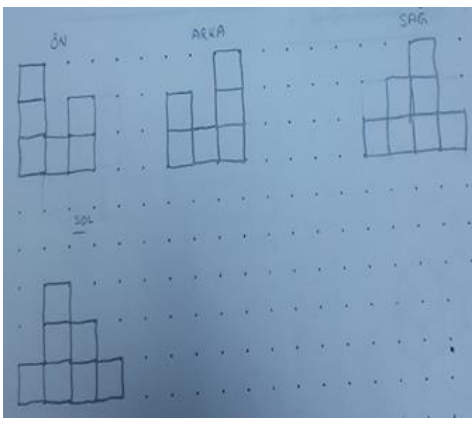
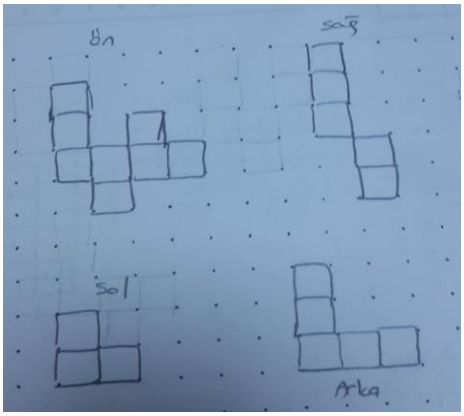
2b. 

Figure 2. Samples of correct drawings by the participants

3a. 

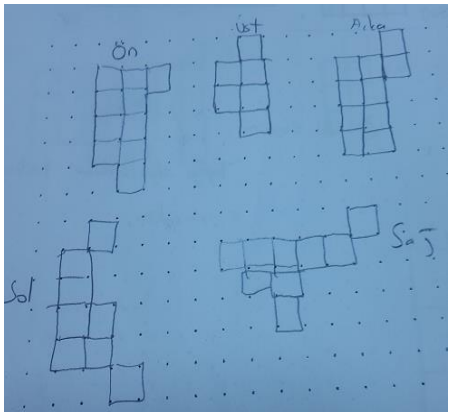
3b. 

Figure 3. Samples of incorrect drawings by the participants

Table 3. *The participants' recognition of the symmetry between the left and right and the frontal and rear views of the structure*

Symmetry	f	%
Recognized the symmetry between the views	12	35
Did not recognize the symmetry between the views	22	65
Total	34	100

Table 3 presents findings about the participants' ability to recognize the symmetry between the right and left and the frontal and rear views of the structure in the viewpoints activity. Table 3 indicates that 35% of the participants were able to recognize the symmetry between the left and right and the frontal and rear views, while 65% were unable to do so. The participants' level of recognizing the symmetry was low.

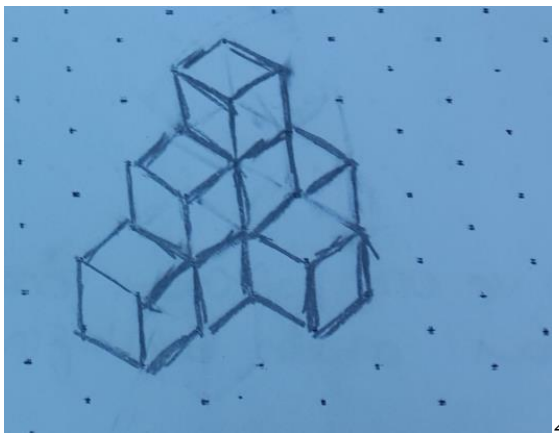
Table 4. *The participants' ability to build a structure with unit blocks using left, right, frontal and rear views of it*

Building the structure	f	%
Able to build the structure with unit blocks	27	79
Unable to build the structure using unit blocks	7	21
Total	34	100

Table 4 presents findings about participants' ability to build a structure with unit blocks using left, right, frontal and rear views of it. The participants were asked to draw four views of structures they built using unit blocks for Table 2, and they were asked to do the reverse for Table 4. Table 4 indicates that 79% of the participants were able to build the structure with unit blocks using left, right, frontal and rear views of it. Only 21% of the participants were unable to build the structure, and participants were more successful at building the structure with unit blocks using four views of it.

Table 5. *The participants' ability to draw a front right view of the structure they built with unit blocks on isometric paper*

Drawing ability	f	%
Could not draw it at all	15	44
Drew it incorrectly	8	24
Drew it correctly	11	32
Total	34	100

4a. 

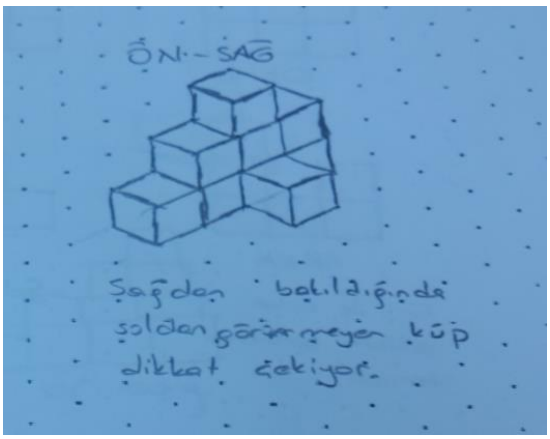
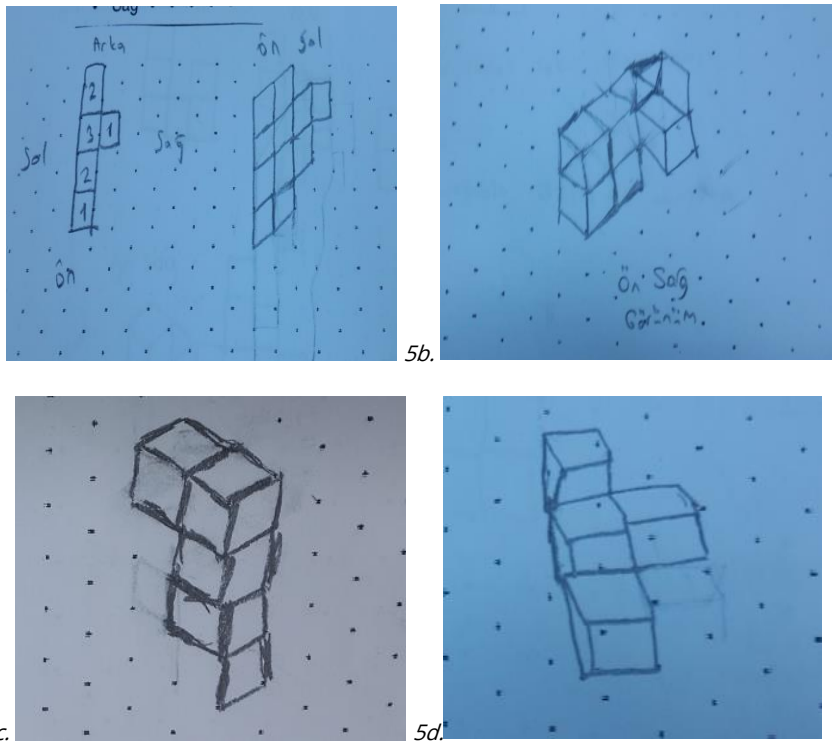
4b. 

Figure 4. Sample correct drawings of front-right view of the structure



5a. 5b. 5c. 5d.

Figure 5. Sample incorrect drawings of the front right view of the structure

Discussion

The geometry sub-field in the mathematics curricula of the first, second, third, fourth and fifth grades) includes learning acquisitions about building three-dimensional structures using unit blocks and drawing them on isometric paper. Students will be able to learn these things by doing activities in classrooms, and homeroom teachers play an important role in teaching these skills to students. For this reason, it is important to know the level of pre-service teachers' skills. This study's results reveal that pre-service teachers' skills in this area are not at the expected level.

The first research question asked the participants to build a structure with unit blocks using an aerial view diagram of it. In general, the participants were successful. The participants were not equally successful at geometrical viewpoint and drawing skills. Less than 50% of the participants succeeded in drawing the front, back, right and left view of the structure. They were most successful at drawing the front view (44%) and least successful at drawing the view from the right (71%). This finding indicates that the participants' ability to depict geometrical viewpoints is not at the desired level.

The right and left and the frontal and rear views of the structures built with unit blocks were symmetrical. Of the participants, 35% recognized this, and 65% did not. This level of recognition is lower than expected. Considering that there were learning acquisitions about the symmetrical connections between structures and symmetry in the second year mathematics curriculum of the NME, it is a subject of discussion how pre-service teachers will be able to teach students about symmetry when they cannot see it themselves.

Of the participants, 79% were able to build a structure with unit blocks using views of it from the right, left, front and back. Thus, the participants were successful at building a structure with unit blocks when they were shown diagrams of it.

The participants were not very successful at drawing the front-right view of the structure they built with unit blocks. Of the participants, 15 were unable to draw it at all, and 8 drew it incorrectly. These results are consistent with those of the study by Duatepe and Paksu (2013), which focused on the drawing skills of pre-service elementary school teachers, and asked them to draw two structures, one made of five identical unit blocks and one made of six identical unit blocks. The study results revealed that approximately one-third of the participants were unable to draw them at all, and only one-fourth of the participants could accurately draw the structure.

Conclusion

Other than pre-service teachers, studies conducted with teachers and students have indicated that the skills of drawing geometrical viewpoints and perspectival drawing, which are directly related to spatial reasoning, is a problematic field in general. Dogan, Temur and Tertemiz (2012) conducted an observational study of elementary school teachers' activities during lessons, and analyzed their thoughts about geometry pedagogy using van Hiele levels. The study results indicated that the teachers taught geometry classes using activities based on their experiences, but did not include enough drawing activities in the lessons (Duatepe & Paksu, 2013). Yolcu and Kurtulus (2010) also conducted a similar study. They both used concrete materials (unit blocks) and computer applications to improve sixth grade students' spatial skills. They also

had students do different activities such as building structures using unit blocks and drawing those structures on isometric paper depicting different views of them. The study results revealed that students improved their visualization skills by building structures with unit blocks and drawing them on isometric paper. These activities showed that geometrical drawing skills are very important for students, teachers and pre-service teachers.

The results of this study and the findings of similar studies suggest that teacher training programs should include activities that use concrete models to improve students' spatial skills and visualization abilities, and, in particular, drawing activities that use isometric paper to improve their drawing skills.

References

Clements, D.H. & Battista, M.T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Eds.), Handbook of research on mathematics teaching and learning (pp. 420-464). Don Mills, ON: Maxwell Macmillan.

Doğan-Temur, Ö. & Tertemiz, N. (2012). İlköğretim birinci kademe öğretmenlerinin geometri öğretimine ilişkin sınıf içi uygulamalarının van Hiele seviyelerine göre irdelenmesi. *DPUJSS, 32*(2), 255-274.

Duatepe Paksu, A. (2013). Sınıf öğretmeni adaylarının geometrik yapılara ilişkin çizim becerilerinin incelenmesi. *Kastamonu Eğitim Dergisi, 21(3),* 827-840.

Hoffer, A. (1981). Geometry is more than proof. *Mathematics Teacher, 74*(1), 11-18.

Ministry of National Education [MoNE]. (2005). İlköğretim ders programları: Matematik programı 1-5 (Elementary curricula programs: mathematics curriculum program for 1-5 grades). Ankara, Turkey: MEB.

Ministry of National Education. [MoNE]. (2009). İlköğretim matematik dersi öğretim programı. Retrieved from: http://ttkb.meb.gov.tr/www/ogretim-programlari/icerik/72.

National Council of Teachers of Mathematics. (2000). Principles and standards for school Mathematics. Reston. VA.

Piaget, J., & Inhelder, B. (1967). *The child's conception of space.* New York, NY: W. W.

Powel, M. J. (1997). Geometry, mathematics in the middle grades: Excerpts from NCTM standards. <http://www.sedl.org/ >. (2005.10.12).

Werthessen, H. (1999). *Instruction in spatial skills and its effect on self-efficacy and achievement in mental rotation and spatial visualization* (Doctoral Dissertation). Columbia University. <http://wwwlib.umi.com/dissertations/fullcit/9634439> (2015.03.24).

Van de Walle, J. A., Karp, K.S., & Bay-Williams, J. M. (2009). *Elementary and middle school mathematics: Teaching developmentally* (7th Edition). *Boston, USA: Allyn & Bacon.*

Yıldırım, A. & Şimşek, H. (2011). *Sosyal bilimlerde nitel araştırma yöntemi.* Seçkin Yayıncılık. Ankara, Turkey.

Yolcu, B. & Kurtuluş, A. (2010). 6. Sınıf öğrencilerinin uzamsal görselleştirme yeteneklerini geliştirme üzerine bir çalışma. *Elementary Education Online, 9*(1), 256-274.